Bachelor's thesis

Information and Communications Technology

2025

Joni Obradovic

Federated brain tumor segmentation with patient-level local differential privacy

TURKU AMK
TURKU UNIVERSITY OF
APPLIED SCIENCES

Bachelor’s Thesis | Abstract

Turku University of Applied Sciences

Information and Communications Technology

2025 | 69 pages

Joni Obradovic

Federated brain tumor segmentation with patient-level local differential privacy

Federated Learning (FL) facilitates collaborative training of machine learning models across distributed datasets held privately by multiple institutions, such as hospitals with sensitive medical imaging data, without centralized sharing of raw patient data.

To further strengthen patient confidentiality, Local Differential Privacy (LDP) introduces controlled, privacy-preserving noise directly into the model's gradients locally at each participating institution, before their aggregation into a global model. This approach is especially relevant in healthcare, where safeguarding patient data is both ethically and legally mandated.

This thesis investigates and implements local differential privacy techniques within a federated learning framework, specifically applied to brain tumor segmentation using medical imaging. The primary goal was to enhance patient-level privacy protections while minimizing degradation in segmentation performance.

A 3D Residual U-Net model was trained under LDP conditions by adding Gaussian noise locally to gradients before central aggregation. Essential differential privacy parameters, including noise magnitude, gradient clipping thresholds, and local training hyperparameters, were explored to achieve a favorable balance between privacy and segmentation accuracy.

Experimental results indicated a slight decrease in segmentation performance, assessed using standard medical imaging metrics (Dice Similarity Coefficient and Hausdorff95 Distance), when incorporating local differential privacy.

However, it was demonstrated that this performance loss could be substantially mitigated by careful tuning of privacy parameters. Ultimately, the study confirms the practical feasibility of local differential privacy in federated medical imaging applications, highlighting that robust patient-level data protection can be effectively implemented with moderate computational overhead and targeted optimization efforts.

Future work could further refine this balance through advanced hyperparameter optimization techniques and alternative noise addition strategies, expanding the potential for secure and privacy-conscious collaborative machine learning in medical imaging contexts.

Keywords:

Federated Learning, Medical Image Segmentation, Local Differential Privacy

Opinnäytetyö (AMK) | Tiivistelmä

Turun ammattikorkeakoulu

Tieto- ja Viestintätekniikka

2025 | 69 sivua

Joni Obradovic

Potilastason lokaali differentiaalinen yksityisyys federoidussa aivokasvainten segmentoinnissa

Federoitu oppiminen (Federated Learning, FL) mahdollistaa koneoppimismallien yhteistoiminnallisen koulutuksen hajautetuissa tietoaineistoissa, joita hallinnoivat itsenäisesti eri organisaatiot, kuten sairaalat, ilman että arkaluonteista potilastietoa tarvitsee jakaa keskitetysti.

Paikallinen differentiaalinen yksityisyys (Local Differential Privacy, LDP) parantaa entisestään potilastason tietosuojaa lisäämällä kontrolloitua, yksityisyyttä suojaavaa kohinaa paikallisesti kunkin osallistuvan organisaation laskemiin malligradientteihin ennen niiden yhteistä aggregointia. Tämä lähestymistapa soveltuu erityisen hyvin terveydenhuollon sovelluksiin, joissa potilastietojen suojaaminen on sekä eettinen että juridinen välttämättömyys.

Tässä opinnäytetyössä tutkittiin ja implementoitiin paikallisen differentiaalisen yksityisyyden menetelmää federoidussa oppimisessa, sovelluskohteena lääketieteellinen kuvantaminen ja erityisesti aivokasvainten segmentointi.

Työn tavoitteena oli vahvistaa potilastason tietosuojaa niin, että segmentointitarkkuus heikkenisi samalla mahdollisimman vähän. Tutkimuksessa käytettiin 3D Residual U-Net -mallia, jonka paikallisiin gradientteihin lisättiin gaussista kohinaa ennen niiden lähettämistä aggregointipalvelimelle. Oleelliset yksityisyysparametrit, kuten kohinan määrä, gradienttien rajausarvo sekä paikalliset koulutuksen hyperparametrit valittiin huolellisesti, jotta saavutettiin suotuisa tasapaino yksityisyyden ja mallin suorituskyvyn välillä.

Tulokset osoittivat, että kohinan lisääminen heikensi hieman segmentointitarkkuutta, jota mitattiin yleisesti käytetyillä kuvantamisen mittareilla (Dice Similarity Coefficient ja Hausdorff95-etäisyys). Samalla kuitenkin havaittiin, että heikkenemistä voitiin merkittävästi vähentää yksityisyysasetuksia optimoimalla.

Tämä työ vahvistaa, että paikallinen differentiaalinen yksityisyys on käytännössä toteutettavissa oleva ratkaisu federoidussa lääketieteellisessä kuvantamisessa ja että vahva potilastason tietosuoja voidaan saavuttaa kohtuullisella laskennallisella lisäkuormituksella ja täsmällisellä hyperparametrien valinnalla.

Jatkossa menetelmää voidaan edelleen kehittää hyödyntämällä edistyneempiä hyperparametrien optimointitekniikoita sekä vaihtoehtoisia kohinanlisäysstrategioita. Tämä laajentaisi turvallisen ja yksityisyystietoisen federoidun koneoppimisen mahdollisuuksia lääketieteellisessä kuvantamisessa.

Asiasanat:

Federoitu oppiminen, lääketieteellisten kuvien segmentointi, paikallinen differentiaalinen yksityisyys

Contents

Figures

Tables

List of abbreviations

(ε, δ)-DP	Epsilon-Delta DP: A relaxed (approximate) version of the ε-DP with small probability of failure, denoted by δ.
$\|\|x\|\|_p$	p-Norm of a vector: A mathematical measure of a vector's length, defined as the p-th root of the sum of the absolute values of its components raised to the power of p. This work primarily refers to ℓ_2-norm (Euclidean norm).
$\triangle_2 f$	Function Sensitivity: Measures how much a function's output changes when its input changes. In the context of DP**,** the required noise is calibrated based on sensitivity, typically using the ℓ_2-norm.
δ	(Delta) Privacy Failure Probability: Represents the small probability that the DP guarantee may not hold. It is typically set as $\delta \leq \frac{1}{N}$ where N is the total number of examples in the dataset.
ε	(Epsilon) Privacy Budget: A measure of privacy loss. Used together with Euler's number e as an exponent in ε -DP. A smaller ε indicates higher privacy, while a larger ε indicates lower privacy.
σ	(Sigma) Noise Scale: Defines the scale of noise added by a noise mechanism. In the context of the Gaussian Mechanism σ represents the standard deviation of the added noise.
AI	Artificial Intelligence: The theory and development of computer systems capable of performing tasks that typically require human intelligence, such as visual perception, speech recognition, decision-making, and language translation.
C	*Norm Bound:* Defines the bound to which model gradients are clipped. Used together with the gradient

	clipping function to ensure that model parameters are constrained within an ℓ_2-norm of C.
DL	Deep Learning*:* A subfield of machine learning (ML) that focuses on neural networks with multiple layers (deep architectures) for complex tasks such as image recognition, natural language processing, and medical imaging.
DP, ε-DP	(Epsilon)-Differential Privacy: A privacy framework where a randomized algorithm ensures that the probability of any fixed response does not change by more than a factor of e^{ε} upon substituting a single individual in the dataset. Also referred to as *"pure DP",* meaning it has no probability of failure.
e	Euler's number: A mathematical constant approximately equal to 2.71828, which serves as the base of the natural logarithm and exponential function.
FL	Federated Learning: A decentralized machine learning (ML) approach where models are trained across multiple users or institutions without requiring them to share their personal data.
LDP	Local Differential Privacy: A privacy mechanism applied on the user's side, ensuring that any released information remains differentially private before being shared.
ML	Machine Learning: A subset of artificial intelligence (AI) that enables computer systems to learn from data and make predictions or decisions without being explicitly programmed.
SGD	Stochastic Gradient Descent: An iterative optimization method for ML/DL training. Unlike gradient descent, which computes updates using the entire dataset, SGD updates model parameters using only a single or a few training examples per iteration.

1 Introduction

Recent trends in healthcare reveal a multifaceted challenge that extends beyond technological limitations. The rising cost of healthcare is driven by several factors, including an aging population that demands more intensive care (Valkonen et al., 2021), escalating administrative expenses, and political pressures that might promote privatization while driving cost cutting measures into public healthcare services (Molinuevo et al., 2017).

Consequently, healthcare professionals face an unsustainable volume of data to be assessed (McDonald et al., 2015), compounded by extensive time spent on reporting and updating multiple patient systems (Vehko et al., 2018). These challenges not only exhaust clinical resources but also complicate collaborative efforts. Furthermore, the effective use and usefulness of the vast amounts of clinical data collected daily are hindered by factors such as bloated administrative decision trees, data privacy concerns, and strict regulatory frameworks like the Medical Device Regulation (MDR) (Union, 2025) and secondary use laws (toisiolaki 552/2019, 2019).

The rapidly evolving field of AI and its potential has introduced Federated Learning (FL) as a promising but somewhat unexplored subfield within Machine Learning (ML) (Antunes et al., 2022; Teo et al., 2024). By enabling distributed training on localized data, FL preserves patient privacy. It does so by eliminating the need to centralize sensitive information. However, even decentralized systems are vulnerable to adversarial privacy breaches during the communication of model updates. Adversaries may exploit the communication channels to extract confidential information (Nasr et al., 2019; Tramèr et al., 2016).

To address these concerns,the objective of this thesis is to investigate the application of Differential Privacy (DP) mechanisms at the local client side of federated learning framework. Local Differential Privacy (LDP) introduces controlled noise into local training of segmentation models, effectively masking individual patient’s contributions of a client dataset while maintaining overall

model utility. By designing and implementing a pipeline that integrates DP settings for the medical imaging domain, this research evaluates the trade-offs between data privacy and model performance.

In the following chapters, we delve deeper into the foundations of FL, medical image segmentation and DP. We detail the methodologies employed, and present experiment results that showcase the privacy vs. utility implications of the LDP approach in segmentation tasks.

2 Federated learning for medical image segmentation

Distributed Learning is a well-studied field that emerges from “learning networks” that were developed as a critique against institutionalized schooling (Illich, 1971), However, FL is quite new, more ML focused version that was first defined in a paper offering Deep Learning approach for decentralized participants (Brendan McMahan et al., 2016) , and it was later refined into a broader definition (Kairouz et al., 2021, p. 4):

> **Federated learning** is a machine learning setting where multiple entities (clients) collaborate in solving a machine learning problem, under the coordination of a central server or service provider. Each client’s raw data is stored locally and not exchanged or transferred; instead, focused updates intended for immediate aggregation are used to achieve the learning objective.

Distinctions between classical ML and FL are illustrated in the following examples (Figure 1, Figure 2) where clients are defined as smartphone users with user’s data.

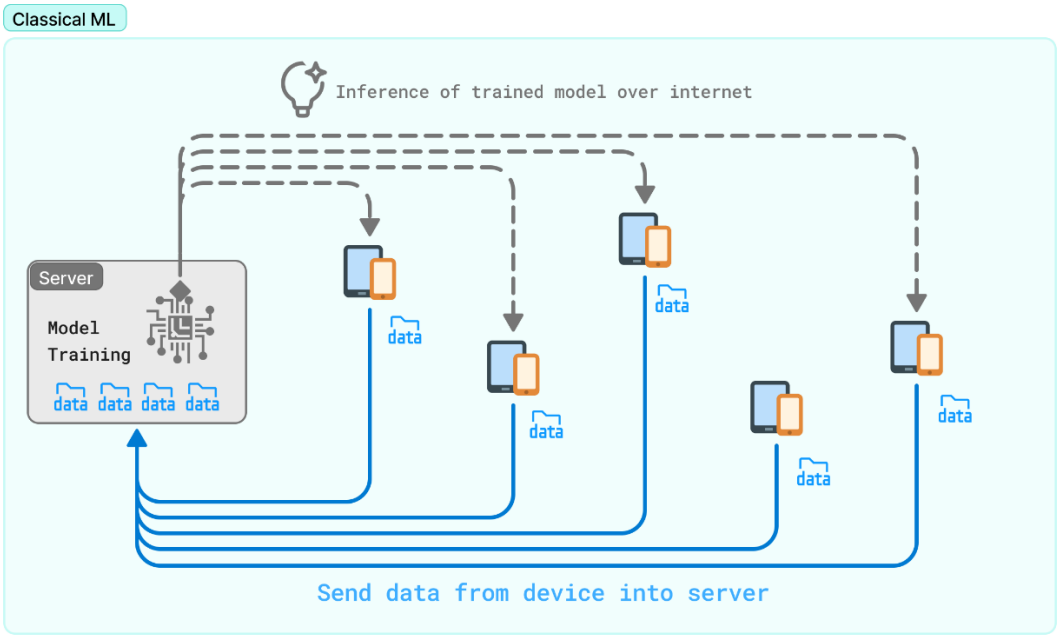

Figure 1. Classical Machine Learning.

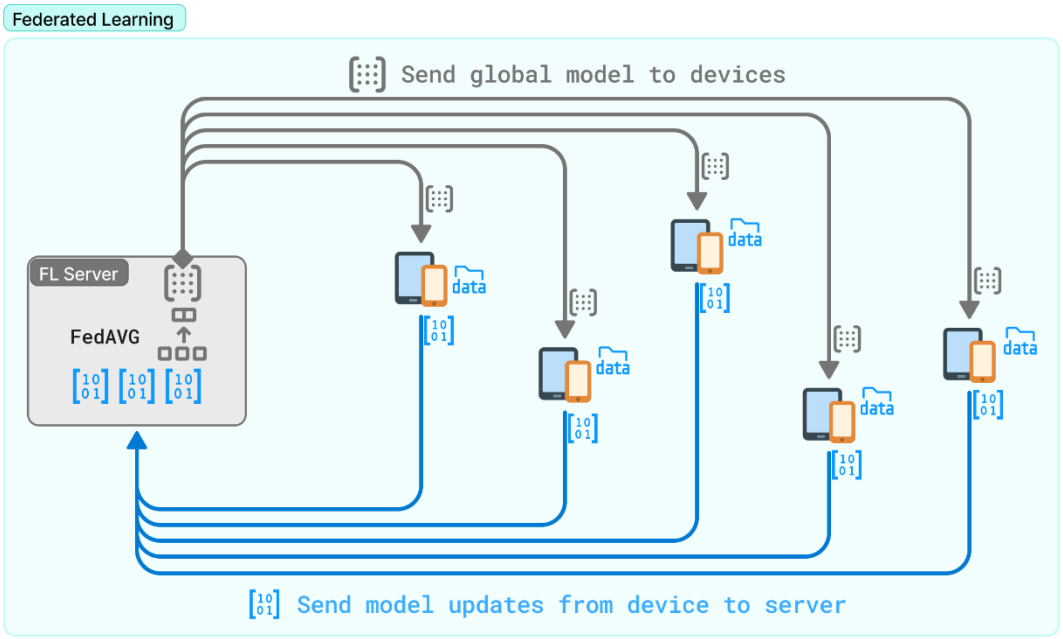

Figure 2. Federated Learning.

Classical ML (Figure 1) and FL (Figure 2) differ fundamentally from in its handling of data and training. Instead of centralizing potentially sensitive data as in classical ML, FL keeps data decentralized and distributes the model parameters for local training updates. This thesis applies the FL paradigm to the task of brain tumor segmentation. The specific implementation used is a refined and modified version of the codebase from the Federated Brain Tumor Segmentation (FeTS) challenge (Pati et al., 2021a). Orchestrated by a central server (Aggregator), this FL setup allows for training large-scale segmentation models across multiple institutions. Crucially, this enhances patient data privacy by keeping information localized while simultaneously benefiting from the diverse datasets available at different sites.

2.1 Cross-silo federated learning

Unlike in cross-device FL where clients are for example mobile devices (Figure 1, Figure 2), this work utilizes cross-silo FL (Huang et al., 2022). Cross-silo FL is a specific topology where participants are organizations or institutions with

substantial computational resources and reliable communication channels. Unlike cross-device FL, involving thousands or even millions of mobile devices, cross-silo FL (Figure 3) features fewer but more powerful clients e.g. hospitals with private collections of patient (in our case MRI) data.

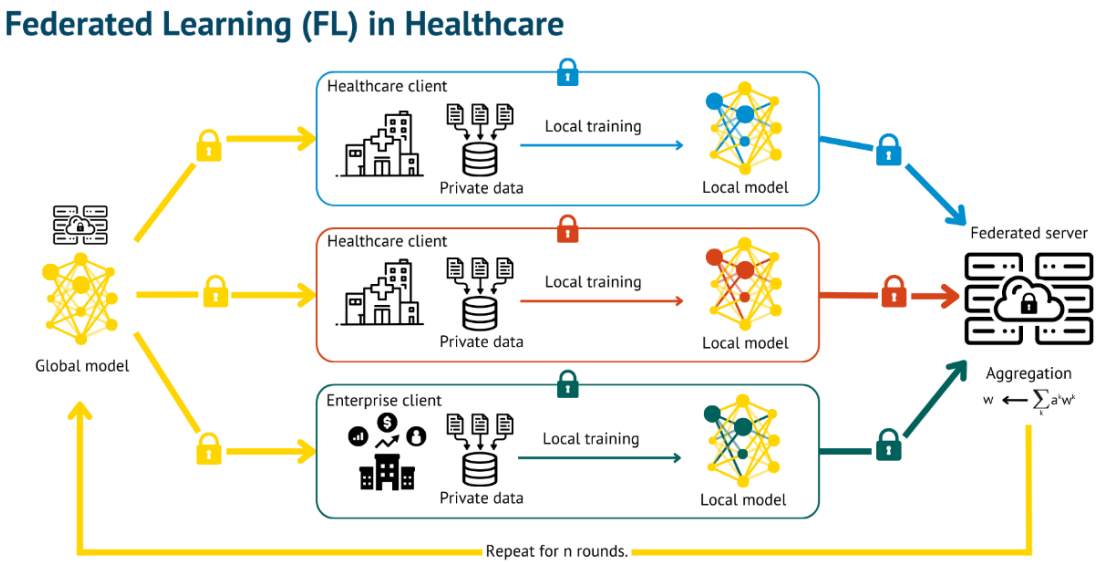

Figure 3. Cross-Silo Federated Learning.

Cross-silo setting is particularly suitable for medical applications because:

- Medical data is subject to strict privacy regulations limiting data sharing between institutions
- Hospitals typically maintain robust computational resources capable of training complex models
- The institutional setting allows for more reliable participation throughout the training process with less likely dropouts

Each client hospital (silo) maintains complete control over its patient MRI data while still contributing to a collaborative brain tumor segmentation model. This preserves patient privacy while enabling access to the statistical power of a larger, more diverse dataset spanning multiple institutions. This approach

allows hospitals to benefit from a collective intelligence without exposing sensitive patient data, addressing both regulatory requirements and ethical considerations in medical data usage.

3 Medical image segmentation

Medical image segmentation is a technique to partition medical images into various regions of interest, thus allowing identification and isolation of anatomical structures or pathological areas. Segmentation serves as a crucial processing step for diagnosis, treatment planning and analysis in clinical applications. Modern Deep Learning techniques have vastly improved to produce sophisticated and accurate segmentations (Hesamian et al. 2019).

3.1 Brain tumor segmentation

Our work relies on professionally processed and annotated open-access multiparametric Magnetic Resonance Imaging (mpMRI (Hao et al., 2011)) dataset (Pati et al., 2021b), consisting of four sequences (Figure 4, Figure 5) that capture different tissue properties (Bitar et al., 2006):

- T1-weighted (T1): Provides clear anatomical detail and gray/white matter distinction
- T2-weighted (T2): Highlights fluids with gray matter appearing brighter than white matter, making it effective for detecting edema and inflammation.
- T1 with contrast enhancement (T1ce): Reveals areas with blood-brain barrier disruption, showing active tumor regions.
- Fluid Attenuated Inversion Recovery (FLAIR): Suppresses cerebrospinal fluid signal while maintaining T2 weighting providing superior contrast for detecting periventricular lesions and subtle abnormalities.

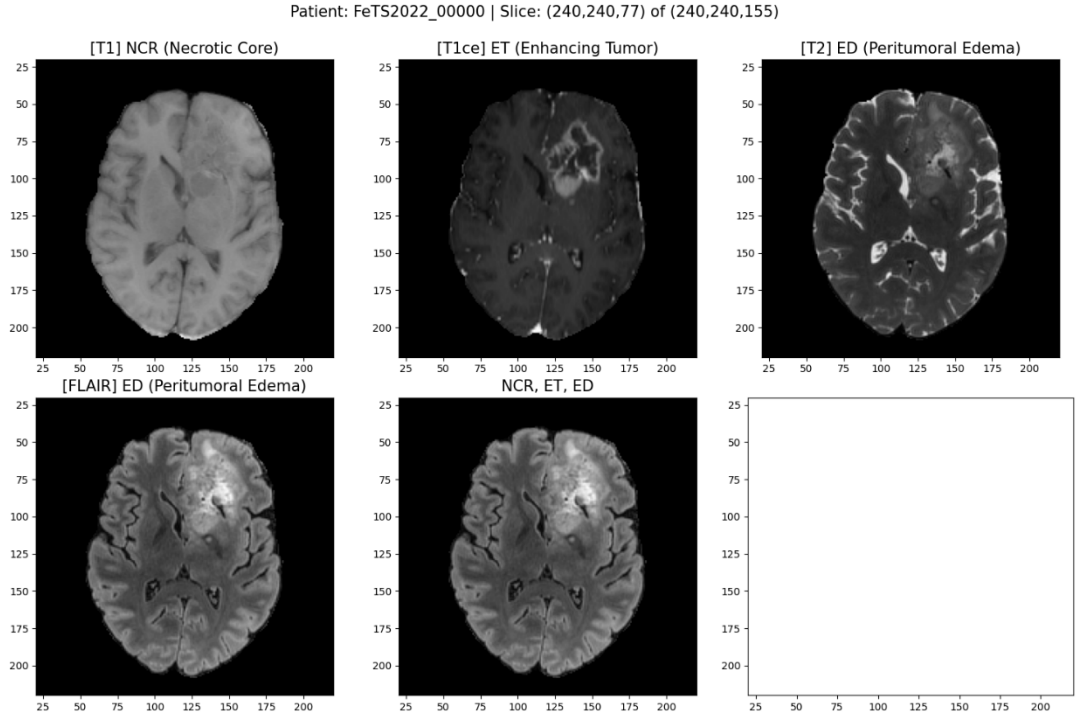

Figure 4. MRI Sequences separately without segmentation mask.

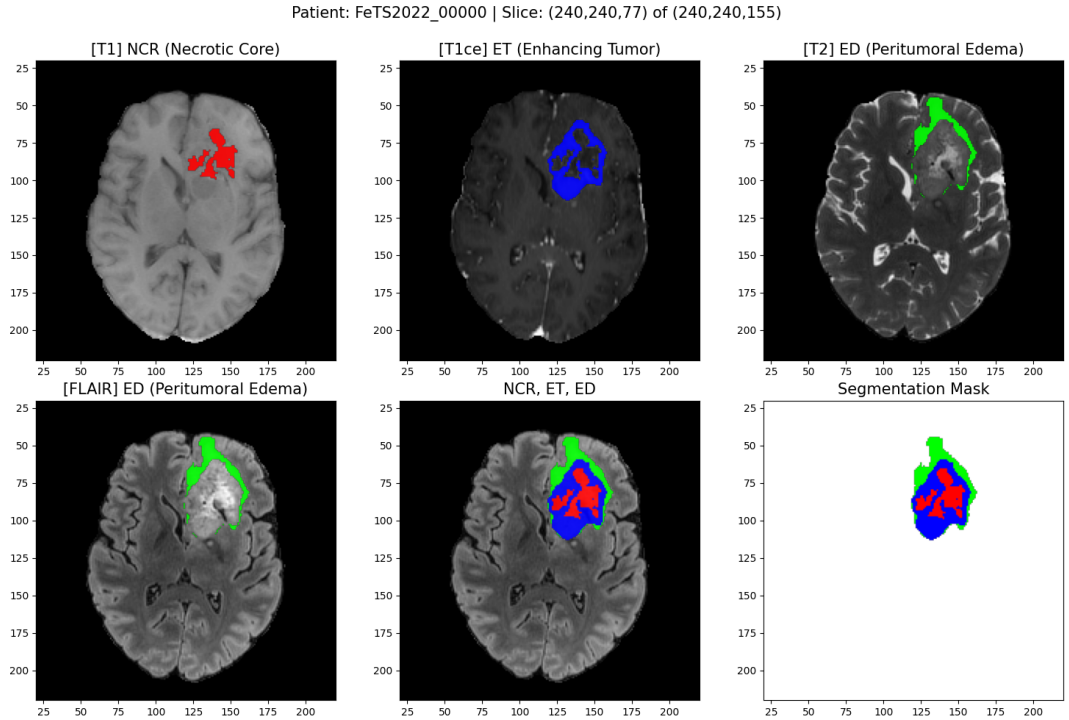

Figure 5. MRI Sequences and segmentation masks with their labels overlayed into each sequence.

These mpMRI datasets are structured as 3D volumes with anisotropic voxel spacing, requiring algorithms capable of processing volumetric data rather than treating each slice independently (Menze et al., 2015).

Because of the nature of tumors and their boundaries, this introduces various challenges (Despotović et al., 2015; D'este et al., 2021):

1. Infiltrative growth patterns: Gliomas often have diffuse boundaries that infiltrate surrounding tissues, making precise delineation difficult even for experienced radiologists
2. Tumor heterogeneity: Different tumor regions (necrotic core, enhancing tumor, peritumoral edema) exhibit different imaging characteristics requiring multi-modal analysis
3. Anatomical variability: Brain structures vary significantly between patients, complicating standardized approaches

The class imbalance problem represents another significant challenge, as tumor regions typically constitute only 1 to 5% of the brain volume. This imbalance can bias segmentation algorithms toward healthy tissue classification, leading to poor sensitivity for tumor detection (Isensee et al., 2021). This is why class penalty weights (λ) are incorporated in our loss function outlined in the following sections. Additionally, volumetric context is critical in accurate segmentation, as analyzing the 3-dimensional spatial relationships between consecutive slices provides crucial information about tumor morphology that would be missed in slice-by-slice processing (G. Wang et al., 2019).

3.2 3D Residual U-Net model architecture

U-Net is a widely used convolutional network for segmentation, originally developed for biomedical images. It has an encoder-decoder design, often described as a contracting path (encoder) and an expanding path (decoder)

arranged in a U-shape (Figure 6), hence the name ”U-Net” (Ronneberger et al., 2015; Weng et al., 2015).

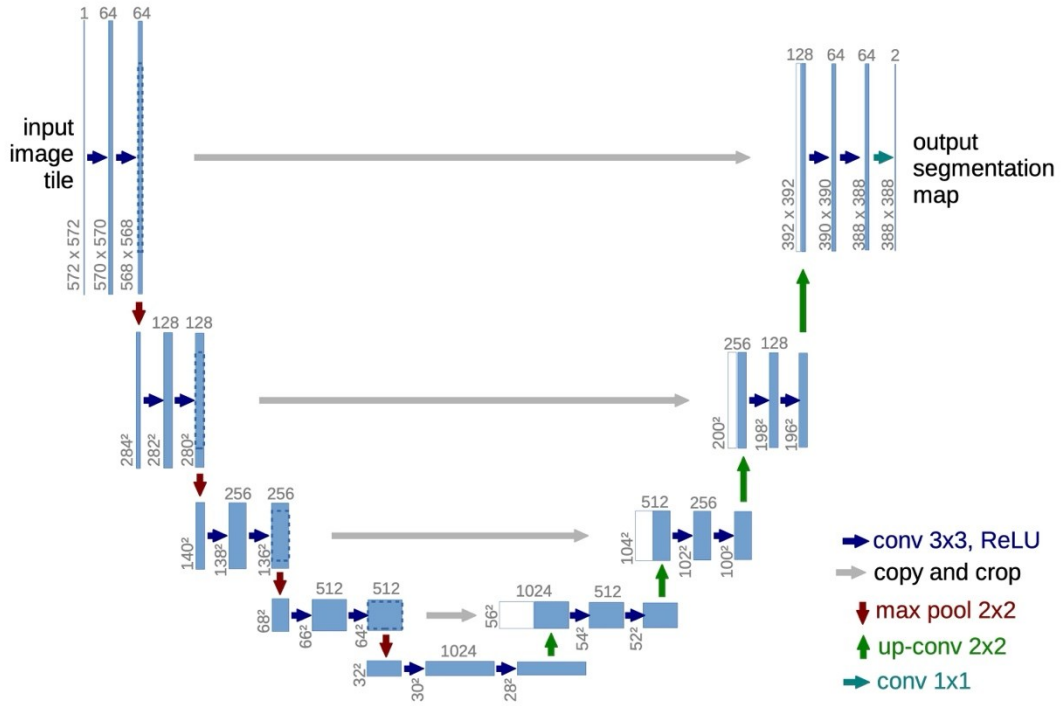

Figure 6. U-Net architecture (Ronneberger et al., 2015).

The encoder block downsamples the input step by step through convolution and pooling layers, capturing contextual features at various scales. The decoder then reverses the process by upsampling and combining these features to enable precise structures in the output segmentation. In a 3D U-Net, these operations are extended to three dimensions, so the network processes volumetric data (e.g. MRI scans) and leverages 3D context across slices (Zhou et al., 2021).

3.2.1 U-Net and skip connections

A key strength of U-Net is the use of skip connections between corresponding encoder and decoder layers. These skips concatenate high-resolution feature maps from the encoder with the upsampled features in the decoder. By merging

low-level details from early layers with high-level semantic information from deeper layers, the network can recover fine details that might have been lost during pooling U-Net: A Comprehensive Guide to Its Architecture and Applications - viso.ai. This is extremely beneficial for brain tumor segmentation, because tumors have irregular shapes and fine boundaries. The skip connections effectively give the decoder direct access to the shallower features (edges, textures) to refine the tumor outline while still using the encoder's global context.

3.2.2 Residual connection

The Residual variant incorporates residual blocks into the U-Net architecture. Residual learning means each convolutional unit doesn't directly output a new feature map but rather adds a small "residual" to its input via a shortcut connection (He et al., 2015). In practice, this introduces identity skip connections within each layer block, allowing the network to learn residual functions (differences) on top of the input of that block. This technique greatly improves the flow of gradients and information through the network. It addresses the "degradation" problem when training very deep networks, making it easier to train a deeper U-Net without performance dropping off (He et al., 2015). In the context of brain tumor segmentation, this results more complex patterns (e.g. subtle texture differences between tumor and healthy tissue) while still maintaining efficiency.

3.2.3 Patch-based training

One practical consideration in 3D segmentation is Graphics Processing Unit (GPU) memory usage. Full brain MRI scans (data used in this work are $240 \times 240 \times 155$ voxels) are too large for GPU memory all at once with high resolution. Instead, a patch-based training strategy is used: the 3D volume is divided into smaller sub-volumes (patches), in our case of size $64 \times 64 \times 64$ voxels, and the network is trained on these patches. This means the model

learns to segment a smaller cube of the image at a time, which fits in GPU memory. The output of N patches is then constructed into original shape by various tiling and aggregation methods.

Brain MRI tumor segmentation benefits greatly from U-Net's design. The encoder's context capabilities helps the model recognize the tumor in relation to anatomical structures (e.g. distinguishing tumor core vs edema in the whole brain volume), while the decoder with skips ensures that the exact tumor boundaries are properly outlined. This is important because tumors can be small or occupy complex regions that require both global context to find them and local detail to accurately segment them.

U-Net was originally demonstrated on biomedical challenges and showed that it can produce sharp, accurate segmentations even with relatively few training images (Ronneberger et al., 2015). The 3D variant further ensures that slice-to-slice continuity in volumetric data is preserved, since it examines a neighborhood in all three dimensions. In brain tumor segmentation challenges (like FeTS), 3D U-Net architectures with residual and attention enhancements have been top performers, highlighting that this architecture is well-suited to the domain(Pati et al., 2021b).

4 Mathematical formulation in the context of federated segmentation tasks

Our FL problem can be formulated as a distributed optimization task where the objective is to find 3D U-Net model parameters that minimize a global loss function across all hospitals' datasets without directly accessing the data.

4.1 Federated optimization task

Let $w \in \mathbb{R}^d$ denote the trainable parameters of our 3D-UNet model for brain tumor segmentation, where $\mathbb{R}$ represents real numbers and d the total number of parameters (dimensions). We define the global objective function $F(w)$ that measures discrepancy between predicted masks (PM) and ground truth (GT) across all hospitals as:

$$F(w) = \sum_{k=1}^{K} \frac{n_k}{n} F_k(w)$$

Where:

K is the number of participating hospitals

n_k is the number of samples (patients MRI volume) at hospital k

$n = \sum_{k=1}^{K} n_k$ is the total number of samples across all hospitals

$F_k(w)$ is the local objective function for hospital k:

$$F_k(w) = \frac{1}{n_k} \sum_{i \in \mathcal{D}_k} L_i(w)$$

Where:

$\mathcal{D}_k$ is the local dataset at hospital k

$L_i(w)$ is the loss function evaluated on the i-th training sample.

4.2 Loss function

For our brain tumor segmentation task, a modified Dice-Sørensen coefficient (DSC) (Sørensen, 1948) is used as the foundation of our loss function(Pati et al., 2021b):

$$\mathrm{DSC} = \frac{2|p \cap y|}{|p| + |\mathrm{y}|}$$

Where:

p is the probability map output by our network's final softmax layer, providing values in range $[0,1]$ for each voxel and class, with probabilities summing to 1 across classes per voxel

y is the one-hot encoded binary segmentation mask (Ground Truth) for multi-class segmentation

The Mean Class Dice (MCD) loss function accounts for class imbalance in tumor sub-regions by computing the Dice coefficient separately for each class. The loss for a training sample i is defined as:

$$L_i(w) = \mathrm{MCD}\ (f(x_i; w), y_i, C, \lambda)$$

Where:

x_i represents the input data

y_i represents the GT segmentation with multi-class labels for different tumor sub-regions

$f(x_i; w)$ represents the model's prediction for input x_i

C is number of classes e.g. background and tumor sub-regions

λ is a vector of class-specific penalty weights

Overall, as shown in the final formula $\mathrm{Dice}_c(p, y)$, this implementation computes the intersection using element-wise multiplication of prediction and ground truth values and normalizes by the sum of p and y elements. The MCD as a whole is computed as:

$$\mathrm{MCD}\ (p, y, C, \lambda) = \frac{1}{C}\sum_{c=1}^{C} \lambda_c \left(1 - \mathrm{Dice}_c(p, y)\right)$$

where $\mathrm{Dice}_c(p, y)$ is the Dice coefficient for class c:

$$\mathrm{Dice}_c(p, y) = \frac{2\sum_i p_{c,i}\, y_{c,i}}{\sum_i p_{c,i} + \sum_i y_{c,i} + \beta}$$

The extra small term β is added to prevent division by zero.

This coefficient is used during evaluation to assess segmentation performance. By optimizing $1 - \mathrm{DSC}$, models are trained to maximize the overlap between predicted segmentations and ground truth.

The class-specific penalty weights λ_c are calculated from the training data to adress class imbalance. For each class c, the penalty weight is computed as:

$$\lambda_c = \frac{N_{total}}{C \times N_c} \times \frac{1}{Z}$$

Where:

N_c is the number of voxels belonging to c in training data

N_{total} is the total number of voxels across all classes

C is the total number of classes

Z is normalization factor to ensure $\sum_{c=1}^{C} \lambda_c = 1$

Penalty weight scheme encourages the model to pay more attention to smaller tumor regions, that would otherwise be overlooked in favor of the prevalent background class, by assigning higher penalties.

4.3 Evaluation metrics

During training and evaluation, healthy tissue and three tumor sub-regions are assessed(Pati et al., 2021b):

0. Healthy Tissue → Label 0
1. Enhancing Tumor (ET) → Label 4
2. Tumor Core (TC) → Label 2 and 4
4. Whole Tumor (WT) → Label 1, 2 and 4

Where 0,1,2,4 are the output label numbers used in metrics.

Segmentation quality is evaluated with:

1. Dice Similarity Coefficient (DSC), the same metric we use in our loss function measures spatial overlap between PM and GT.

2. 95th percentile Hausdorff Distance (H95) (Hausdorff, 1914), which computes the max distance between PM and GT boundaries:

$$H_{95}(PM, GT) = \max\left\{\mathrm{Perc}_{95}\{d(p, GT): p \in PM\}, \mathrm{Perc}_{95}\{d(g, PM): g \in GT\}\right\}$$

where $d(x, Y) = \min_{y \in Y} \|x - y\|$ is the distance of point x to set Y.

H95 is more sensitive to local differences while DSC gives more global measure of the overlap. This comes with a cost of H95 being computationally more demanding.

4.4 Federated averaging- and client selection-algorithm

Initially, FL implementations used Federated Stochastic Gradient Descent (FedSGD), which requires the clients to compute gradients on local data and send model updates after processing single batches. While straightforward, FedSGD demands excessive communication between clients and the central server. To address these issues, Federated Averaging (FedAVG) was introduced (Brendan McMahan et al., 2016) allowing participants to train multiple local epochs before sending the trained model parameters. This approach significantly reduces communication overhead while enabling model convergence.

In our implementation, rather than using the standard FedAVG, we utilize two specific algorithms for client selection and model aggregation designed to handle challenges like data heterogeneity (non-IID data) and optimize the FL process (Khan et al., 2023, 2024). These methods aim to improve upon the standard FedAVG approach by incorporating more sophisticated techniques for choosing which clients participate in each round and how their contributions are combined.

4.4.1 Client selection

For selecting clients in each communication round, we adopt the strategy proposed by Khan et al., 2023. This method aims to ensure fair participation while managing the communication load. Instead of involving all clients in every round, a subset (e.g., 20%) is chosen. The selection process works as follows:

1. Randomization: The list of all available clients is initially randomized.
2. Sliding Window: A sliding window moves across this randomized list to select the subset of clients for the current round. For instance, if 33 clients exist and 20% (6 clients using floor division) are selected per round, round 1 selects clients at indices 0-5, round 2 selects indices 6-11, and so on.

3. Reshuffling: Once the sliding window has traversed the entire randomized list, ensuring each client has participated roughly the same number of times, the list is reshuffled, and the process repeats.

This approach ensures that while only a fraction of clients participate in any single round (reducing communication), all clients contribute to the global model over time in a non-deterministic fashion, preventing the same group of clients from being selected repeatedly in consecutive rounds (Khan et al., 2023).

4.4.2 Harmonic Similarity weighted Aggregation

For the aggregation of model parameters at the server, we employ the Harmonic Similarity Weighted Aggregation (HSimAgg) algorithm, an advancement of the SimAgg method specifically designed for robustness in FL settings like brain tumor segmentation (Khan et al., 2024). A key challenge in FL, especially with non-IID data typical in medical imaging, is the potential divergence of model parameters sent by different clients. HSimAgg (Figure 7) addresses this by using a weighted aggregation scheme based on similarity and sample size, incorporating the harmonic mean for the final aggregation step to handle outliers effectively.

Algorithm 1 HSimAgg aggregation algorithm

procedure WEIGHT AGGREGATION(C^r, p_{C^r})
  $\epsilon \leftarrow 1e-5$ ▷ C^r = set of collaborators (at round r)
  $\hat{p}$ = average(p_{C^r}) using **Eq. 1** ▷ p_{C^r} = parameters of the collaborators in C^r
  for c in C^r **do**
    Compute similarity weights u_c using **Eqs. 2** and **3**
    Compute sample weights v_c using **Eq. 4**
  for c in C^r **do**
    Compute aggregation weights w_c using **Eq. 5**
  Compute master model parameters p^m using **Eq. 6**
  return p^m

Figure 7. HSimAgg aggregation algorithm (Khan et al., 2024).

The HSimAgg process at the server for a given round involving a set of participating clients is as follows:

1. Calculate average parameters: The server first computes the unweighted average of the parameters received from all participating clients.
2. Calculate similarity: The inverse distance (similarity) of each client parameters from the average is calculated. This measures how close each client's update is to the central tendency.
3. Normalize similarity weights: These similarities are normalized to obtain similarity weights for each client. Clients closer to the average receive higher weights.
4. Calculate sample size weights: To account for varying amounts of data at each client, sample size weights are calculated based on the number of data samples per client.
5. Combine weights: The similarity weights and sample size weights are combined to form the initial aggregation weights for each client.
6. Harmonic mean aggregation: Finally, the global model parameters for the next round are computed using these weights and the received parameters. HSimAgg utilizes the harmonic mean in its aggregation formula, which is particularly effective at mitigating the influence of extreme values or outliers (Khan et al., 2024).

The aggregated model is then dispatched back to the next set of selected clients for the subsequent training round. By using HSimAgg, the aggregation process becomes more robust to diverging updates from heterogeneous clients, aiming for a more stable and accurate global model (Khan et al., 2024).

5 ε-Differential privacy

Differential Privacy is a mathematical privacy framework that provides guarantees on how much information can leak from any individual in a dataset. The most known definition of Differential privacy was given in *Algorithmic Foundations of Differential Privacy* (Dwork et al., 2014, p. 5):

> "Differential privacy" describes a promise, made by a data holder, or curator, to a data subject: "You will not be affected, adversely or otherwise, by allowing your data to be used in any study or analysis, no matter what other studies, data sets, or information sources, are available."

This principle was formulated in response to earlier work on privacy (Dalenius, 1977), which was deemed too strict for practical applications.

ε-DP was introduced by Cynthia Dwork (Dwork, 2006) and since then, ε-DP has become the most common approach for data privacy in research and commercial applications. Notably, it has been adopted in real-world systems such as the 2020 U.S. Census (Abowd et al., 2022) and Apple (*Learning with Privacy at Scale Differential Privacy Team, Apple*, n.d.). From now on, when we talk about differential privacy as DP, it refers to ε-DP.

As outlined earlier, FL enables training models across distributed hospitals or devices without centralizing patient data. However, even when only sharing model updates, one can leak sensitive patient information via advanced attacks such as model inversion or membership inference (Nasr et al., 2019; Tramèr et al., 2016), in which DP comes into play.

DP provides an additional layer of protection by injecting calibrated noise to the training process, making it mathematically infeasible for an attacker to meaningfully construct any individual patient's information from the model parameters (Dwork, 2006). This is crucial in domains like medical imaging, where data are highly sensitive and subject to strict privacy regulations. By ensuring that outcomes (e.g. the trained model parameters) are insensitive to any single patient's data, differential privacy helps address ethical and legal

concerns while still allowing collaborative cross-silo learning on medical datasets.

The following sections outline the formal definition and key concepts of DP, and how they apply to implementing our patient-level local differential privacy in a federated brain tumor segmentation problem.

5.1 Mathematical framework of differential privacy

Formally, DP is defined with respect to adjacent datasets – typically, two datasets that differ of a single individual. For our approach, a single individual is defined as single patient's MRI sequences where we extract the training patches from.

Let D and D' be two adjacent datasets (e.g. one contains a particular patient's sequences and the other does not). A randomized algorithm (Mechanism) $M: D \rightarrow R$ is said to satisfy ε-differential privacy if for all such adjacent pairs and for *all* possible output events $S \subseteq R$:

$$\Pr[M(D) \in S] \leq e^{\varepsilon} \Pr[M(D') \in S]$$

This bounds the amount how much the probability of an outcome can change by the inclusion or removal of one individual's data (Dwork, 2006). The privacy parameter ε controls this bound – smaller ε means the two probabilities are nearly the same, implying stronger privacy. This indicates that even if an adversary with an access to the output of M, which in our case would be the trained model parameters, cannot confidently extract information whether any specific patient's data was used in the training or not.

Essentially e^{ε} gives the maximum factor that the output distribution changes and this is why ε is called "privacy budget" – smaller budget assumes higher privacy but at the cost of higher noise and reduced accuracy, which is further discussed next.

5.1.1 Privacy–utility trade-off

A fundamental aspect of DP is the tradeoff between privacy and utility. To satisfy the higher (small ε) DP levels, the mechanism has to inject more noise, which degrades the accuracy (utility) of the model. Conversely, a larger ε permits more accurate outputs but weaker privacy guarantees. If $\varepsilon = 0$, any input data will yield identical outputs which naturally makes model training impossible. Thus, selecting appropriate ε (or σ and δ which are discussed later) requires balancing based on the type of model, requirements and goals of the underlying application or its regulative entities. In medical FL, for example, one might choose a moderate ε that provides meaningful patient privacy without rendering the trained model diagnostically useless. This balance should be eplored through examination of regulation and guidelines to conclude acceptable privacy budget and the implicated risks.

5.2 Sensitivity and additive noise mechanisms

As mentioned previously, most DP mechanisms achieve privacy by adding noise calibrated to the sensitivity of the query or computation. The sensitivity of a function f is the maximum change in functions output caused by changing a single individual's data. For training of our segmentation model, we define our sensitivity as:

$$\triangle f = \max_{w,w'} \|f(w) - f(w')\|_p$$

Where w, w' are trained model parameters with data differing a single patient and p is the norm, either ℓ_1 for pure-DP or (in our solution), ℓ_2 (Weisstein).

Bounding this sensitivity is necessary to avoid possibly infinite sensitivity because in theory, model gradients can have infinite values. Bounding the model gradients can be achieved with method called gradient clipping (Abadi et al., 2016; Pascanu et al., 2012):

$$\bar{g} \leftarrow g \div \max\left(1, \frac{\|g\|_p}{C}\right)$$

Where

$\bar{g}$ is processed gradient

g is the calculated gradient

C is the clipping bound

When single patient's data is processed through the model, we take the calculated gradient g and take its p-norm. If $\|g\|_p > C$, we scale g into C. If $\|g\|_p \leq C$, the g is kept unchanged. This guarantees that w, w' have maximum change of C and our sensitivity is bounded.

Once $\triangle f$ is calculated, noise can be added proportionally to the sensitivity. The two most common noise-addition mechanisms are Laplace- and Gaussian mechanism.

5.2.1 Laplace mechanism

In Laplace mechanism, noise sample is drawn from the Laplace distribution:

$$M_{Lap}(D) = f(D) + Lap(0, b)^d$$

Where:

$$b = \frac{\triangle f}{\varepsilon}$$

And d is the dimension of f.

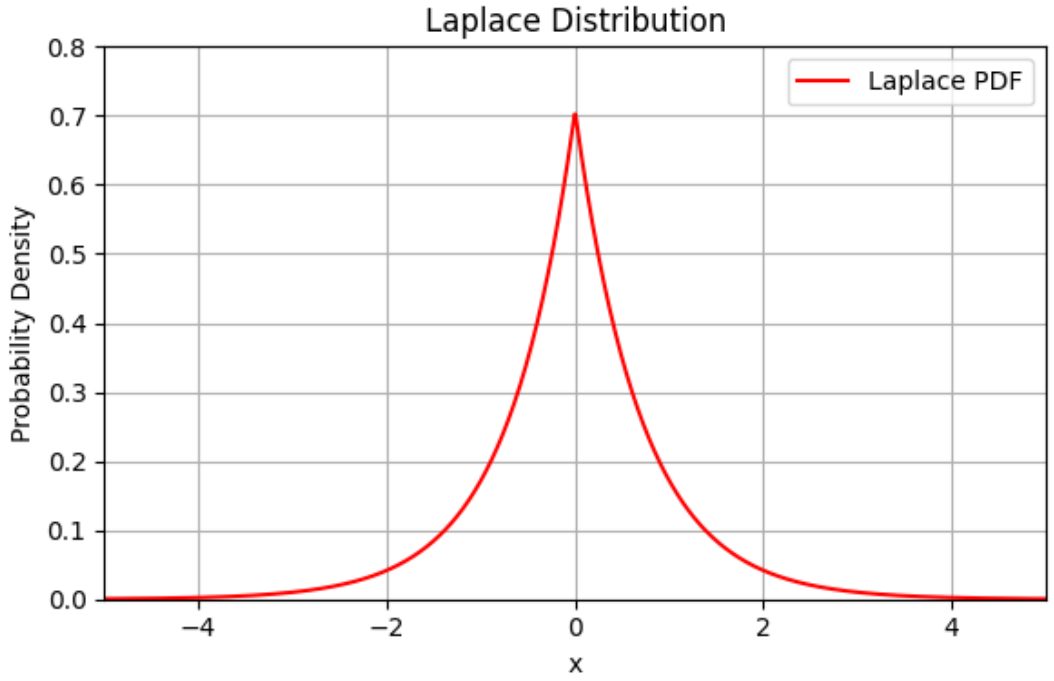

Figure 8. Laplace Distribution.

This (two-sided exponential) mechanism (Figure 8) satisfies pure ε-DP for numeric queries. The Laplace mechanism is straightforward and provides ε-DP for queries with bounded ℓ_1-sensitivity. It is often used for releasing aggregated statistics like counts, sums, or histograms.

5.2.2 Gaussian mechanism

The Noise sample is drawn from a Gaussian (normal) distribution (Figure 9) with mean 0 and standard deviation σ proportional to the ℓ_2-sensitivity.

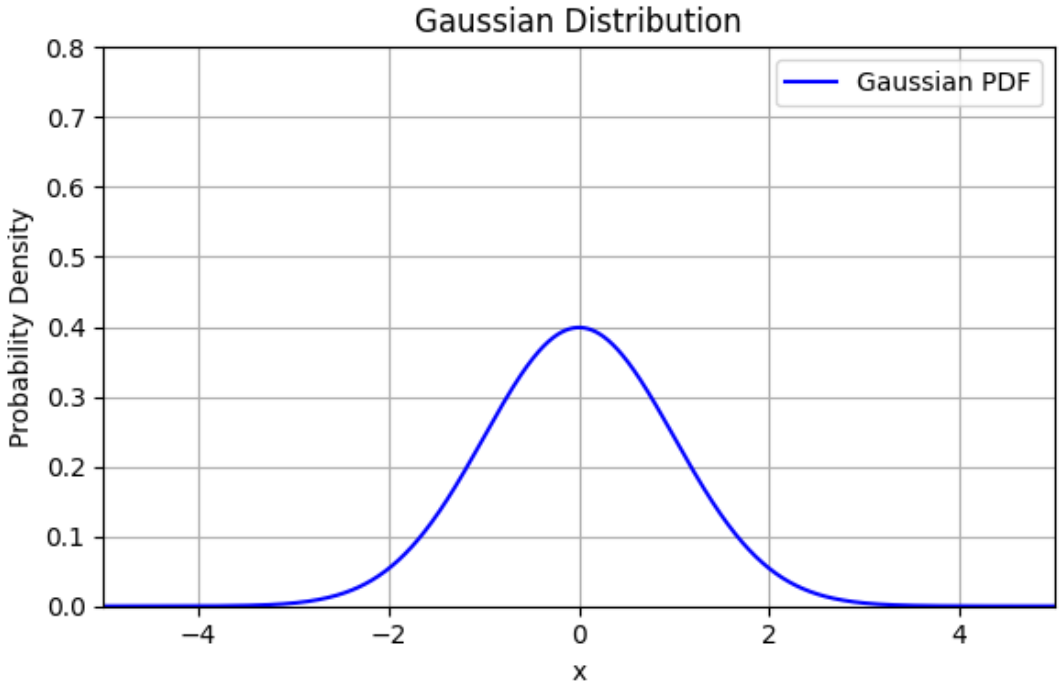

Figure 9. Gaussian Distribution.

Because Gaussian noise has a non-zero probability of very large deviations, it cannot satisfy pure DP – instead, it achieves approximate-*DP* (discussed in next section) by allowing a small probability δ of the privacy bound being violated.

$$M_{Gau}(D) = f(D) + \mathcal{N}(0, \sigma^2)^d$$

Where:

$$\sigma = \frac{\triangle f \sqrt{2 \ln\left(\frac{1.25}{\delta}\right)}}{\varepsilon}$$

The Gaussian mechanism is widely used in machine learning because many vector-valued computations have naturally bounded ℓ_2-norm, and Gaussian noise in high dimensions tends to be mathematically easier to control (Dwork et al., 2014).

In both mechanisms, the noise is calibrated to the worst-case influence of any single individual (via the sensitivity) so that the presence or absence of that

individual’s data has limited impact on the noisy output. The privacy guarantee is maintained regardless of an adversary’s background knowledge and the privacy level is adjusted with privacy-budget parameter (Abadi et al., 2016).

6 From pure to approximate differential privacy

In classical (pure) differential privacy where $\delta = 0$, the privacy guarantee holds for every possible outcome. However, this can lead to unreasonable noise scales for complex tasks like deep learning; it often requires adding so much noise that the accuracy of the model suffers beyond the desired functionality. This usually derives from the fact that Laplace mechanism requires the use of ℓ_1 sensitivity which scales linearly and results massive noise scales in multi-dimensional cases (Figure 10) where ℓ_1-sensitivity grows orders of magnitude faster compared to ℓ_2.

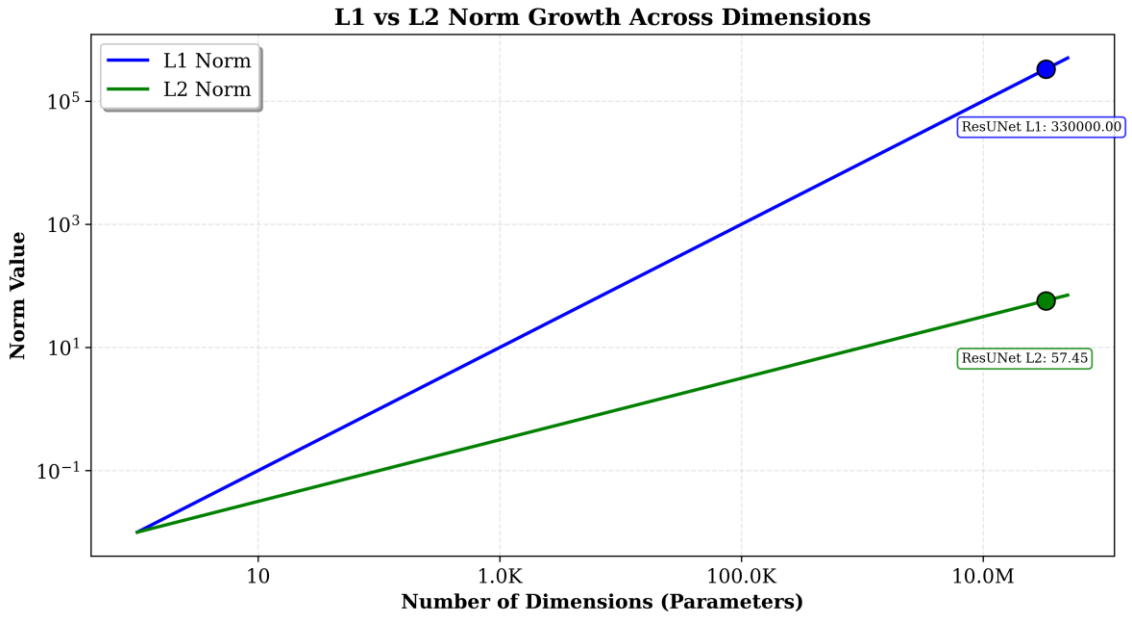

Figure 10. ℓ_1- and ℓ_2-norm scaling on multi dimensional setting.

This can be further visualized with models like the ResNet-50, which has millions of parameters, by feeding 2 different datasamples into the network and outputting their norm-distributions and -statistics:

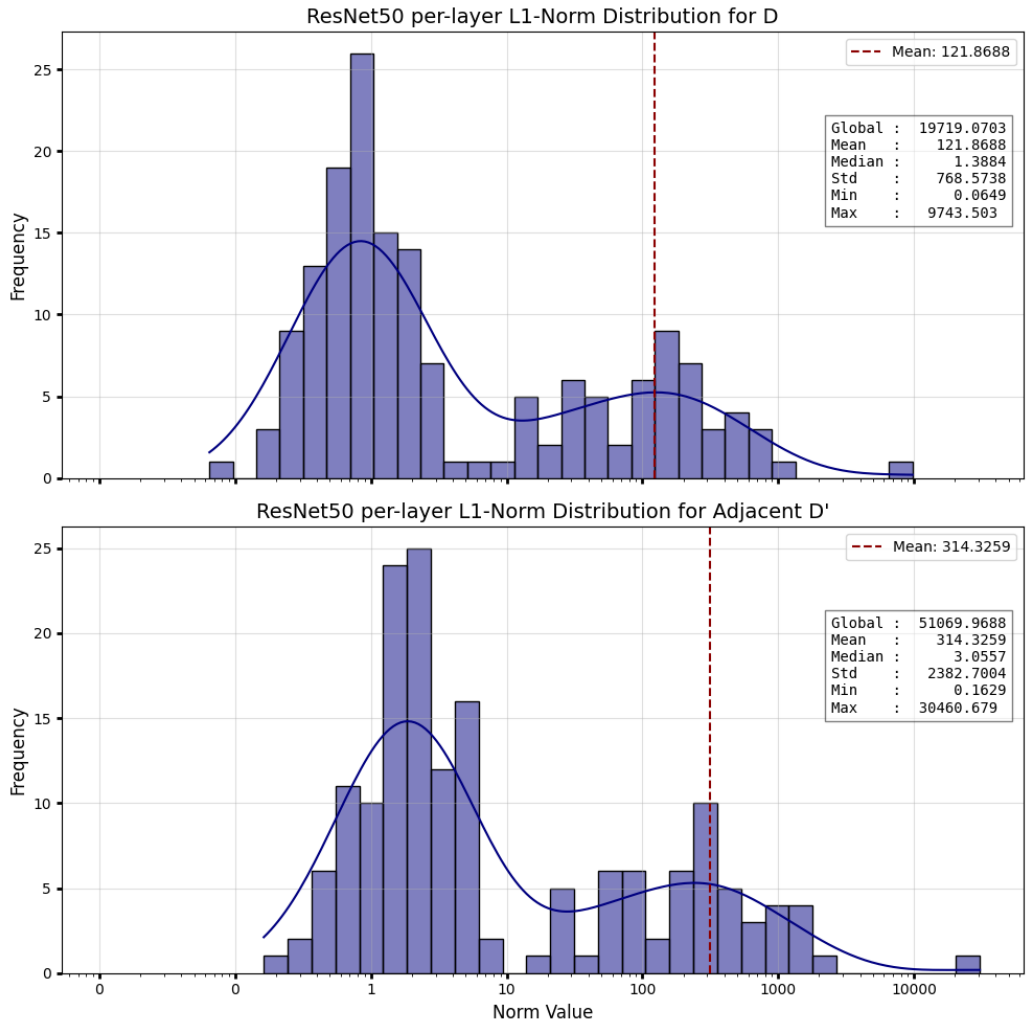

Figure 11. Gradient ℓ_1-norms for ResNet50 between D and D'.

Inspecting ResNet50's gradient data (Figure 11) we can see that the means of per-layer ℓ_1-norms for D and D' are 121 and 314, respectively whereas the corresponding mean ℓ_2-norms (Figure 12) are only 0.24 and 0.46.

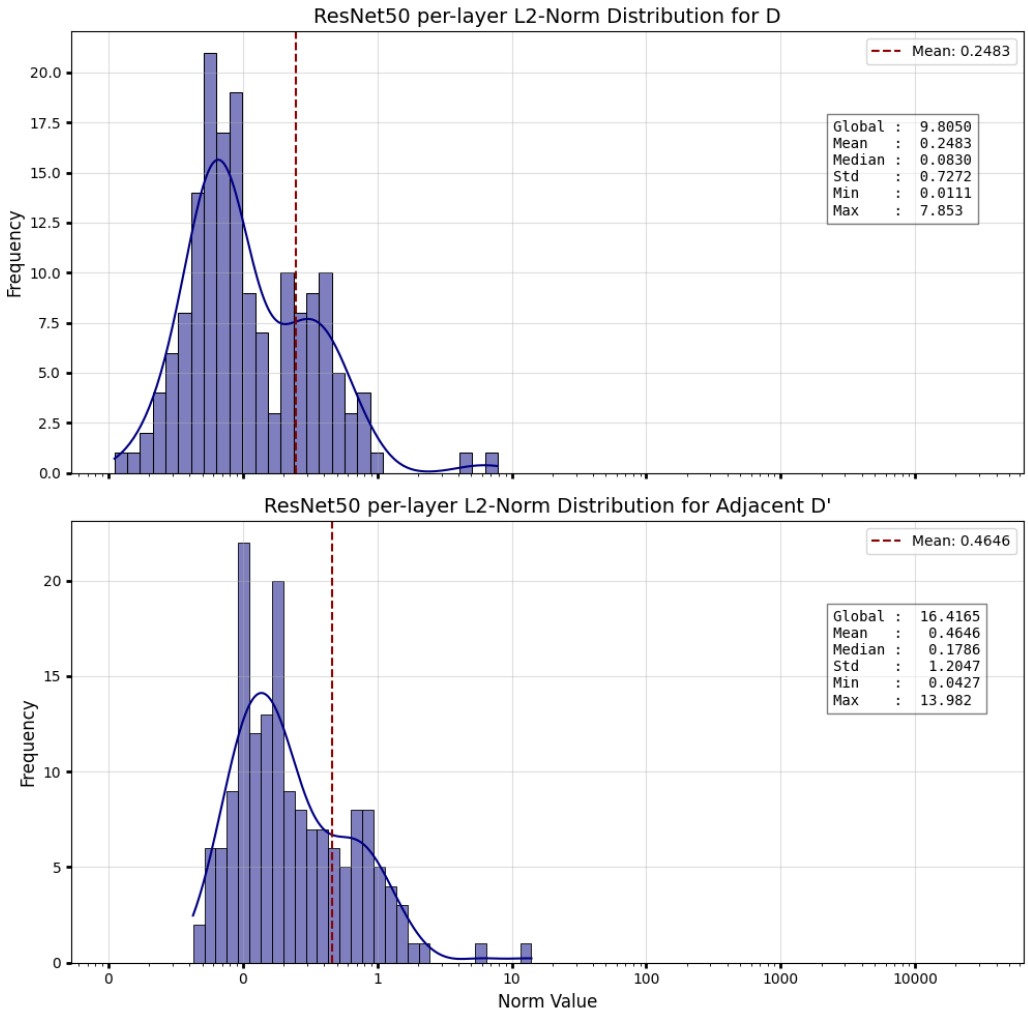

Figure 12. Gradient ℓ_2-norms for ResNet50 between D and D'.

In our approach where we calculate norm across the whole model as single vector, the difference gets even larger with global ℓ_1 being 19719 and 51069 while ℓ_2 is only 9.8 and 16.4 (Figure 11, Figure 12). This difference makes pure-DP with Laplace mechanism almost always infeasible for multi-dimensional problems and for that, approximate gaussian mechanism with ℓ_2-norm is commonly used.

Approximate differential privacy introduces a tiny probability of failure in the guarantee via a parameter δ, yielding an (ε, δ)-DP definition.

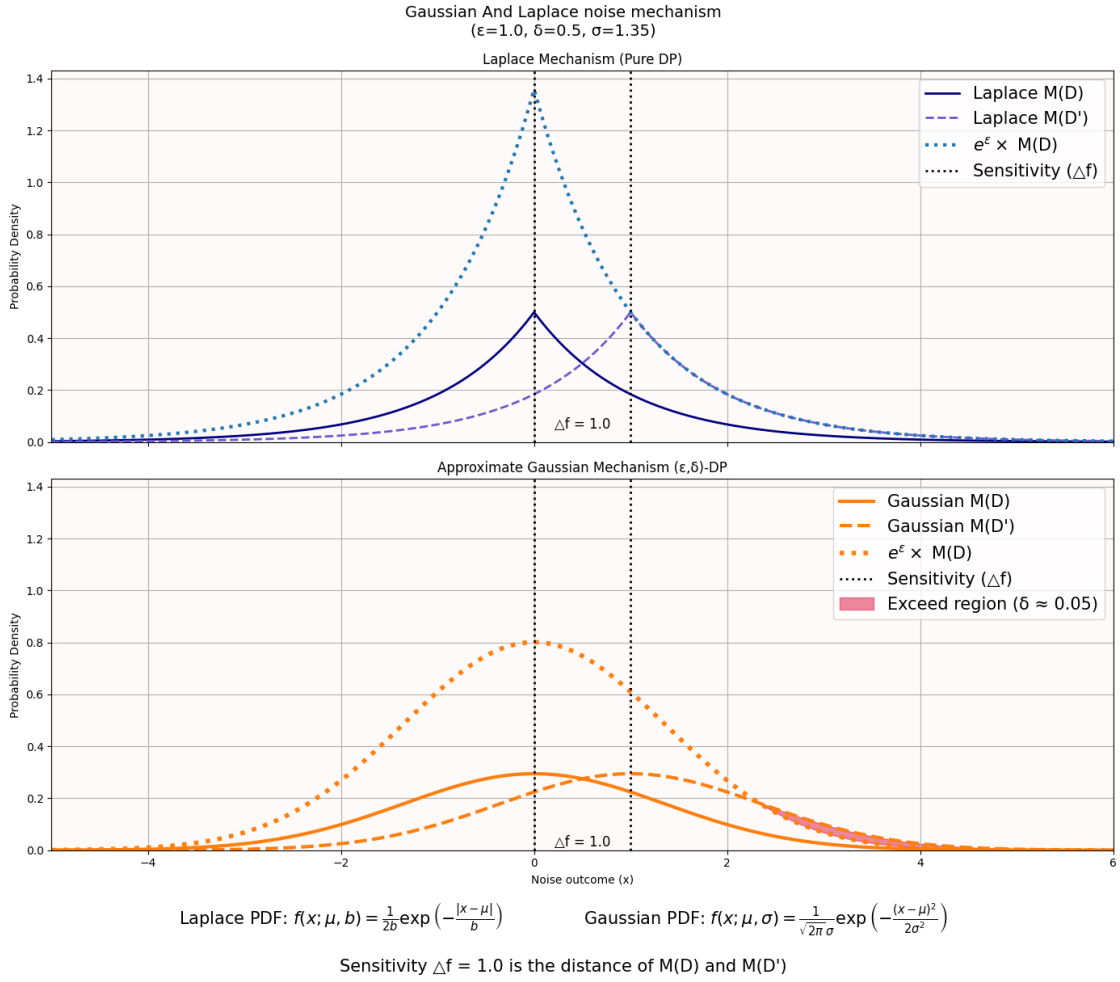

Figure 13. Pure (Laplace) and approximate (Gaussian) noise mechanisms and implication of δ.

Figure 13 showcases the implications of the failure probability which for the sake of clarity, is set really high (0,5) when in reality it is set usually closer to $1E-5$ or $\frac{1}{N}$ where N is the dataset size.

An algorithm satisfying (ε, δ)-DP guarantees that outcomes will shield the participation of single individual by the factor of e^{ε}, except with probability δ (Dwork et al., 2014). The formulation apart from δ, is identical with pure DP:

$$\Pr[M(D) \in S] \leq e^{\varepsilon} \Pr[M(D') \in S] + \delta$$

The benefit of allowing a small δ often permits significantly better accuracy for the same ε, compared to demanding $\delta = 0$. In deep learning FL, this relaxation

is useful because it enables the use of the Gaussian mechanism with ℓ_2 norm, making it feasible to train complex models with somewhat meaningful privacy budgets.

Therefore, (ε, δ)-DP strikes a balance between model performance and patient privacy: the model can learn useful patterns from MRI data while still providing a meaningful privacy guarantee for each patient.

6.1 Local vs. global differential privacy

There are few notable stages in FL pipeline where DP can be applied leading to local and global DP approaches. In Global DP (or server-side DP), the aggregation server collects model parameters from each hospital and then applies noise at the aggregate level. For Local Differential Privacy (LDP), each hospital adds noise to their model updates (or data) before sending anything outside of their private environment. Following illustration (Figure 14, Figure 15) shows the difference between global and local approaches:

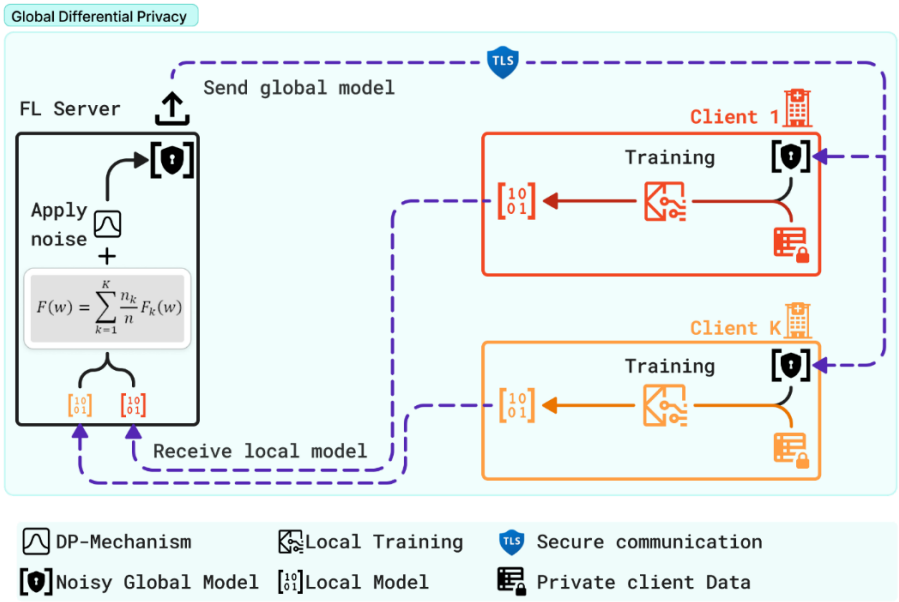

Figure 14. Global differential privacy.

Global-DP (Figure 14) assumes the aggregator is trusted to handle the raw contributions securely until noise is added. Global DP can achieve the same privacy guarantee at potentially lower cost to model utility, because noise is added to an aggregate that includes many users' data.

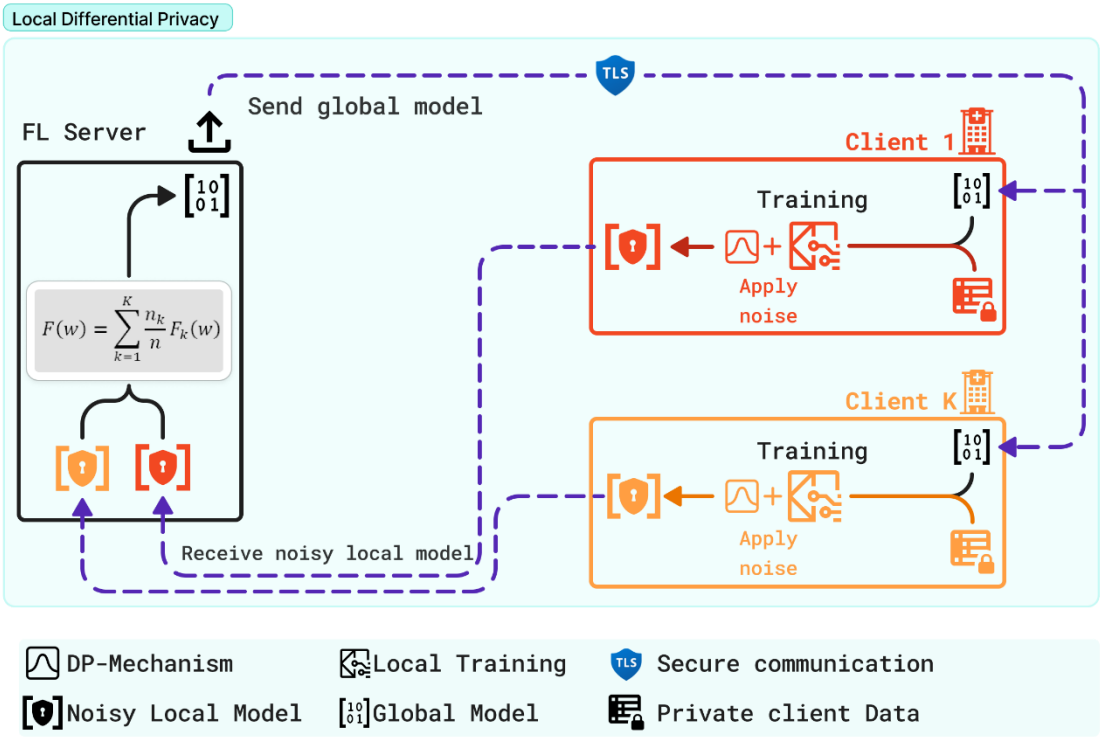

Figure 15. Local differential privacy.

With LDP (Figure 15) no raw information ever leaves the local site – even a potentially untrusted aggregator only sees noisy data. LDP guarantees per-user privacy without requiring any trusted aggregator, which is appealing for sensitive medical data. However, LDP typically demands adding a substantial amount of noise to each user's contribution, since the noise can't be averaged out across users, often resulting in a larger impact on accuracy for a given privacy level (Shan et al.).

Like before, the choices regarding type of DP always involve a trade-off: LDP offers stronger protection against a curious or malicious server, whereas global DP, with a semi-trusted server, can often maintain higher segmentation accuracy for the same privacy budget. In practice, many medical FL systems

lean toward global (server-side) differential privacy with client-level guarantees, because the healthcare clients are by default a lot more trustworthy than, for example, mobile phones by a random person.

7 Privacy accounting in iterative learning

In iterative learning (e.g. training a model with multiple iterations and epochs), we apply a differentially private mechanism repeatedly on the same sensitive dataset. Each application (such as one epoch or one gradient noising in DP-SGD) of the mechanism consumes one unit of ε. DP has a composition property, meaning that combining k DP mechanisms results in an overall privacy loss that accumulates from each application of the mechanism.

Basic composition is the simplest accounting method: if each mechanism independently provides ε-DP, then executing all k mechanisms together provides roughly $\left(\sum_{i=1}^{k} \varepsilon_i\,, 0\right)$-DP. In the common case where each step has the same privacy parameters ε, basic composition says the total privacy cost grows linearly with k, which can also roughly be applied from pure to approximate $(k\varepsilon, k\delta)$ (Dwork, 2006). This linear growth can make the total budget very large after many iterations, which is undesirable – it means weaker privacy guarantees if we naively sum up the losses.

Another term that needs small introduction is Moments (Abadi et al., 2016). It refers to statistical measures that describe the shape and behavior of the privacy loss distribution—a probability distribution that quantifies how much information any single data point leaks through the execution of a DP mechanism. The k-*th* moment of a distribution captures the expected value of the privacy loss raised to the power of k, which helps in tracking how privacy loss accumulates over multiple applications of the mechanism. By leveraging higher-order moments, it allows for more precise tracking of cumulative privacy loss, leading to significantly improved and accurate bounds compared to traditional worst-case bounds. This is particularly useful in iterative learning processes, where privacy loss compounds over multiple training steps and where basic composition methods would result in overly pessimistic estimates.

7.1 Advanced composition

To this privacy budget from increasing unreasonably, researchers have developed advanced composition theorems. By allowing a small δ in the privacy guarantee, one can achieve sublinear growth of ε with the number of compositions. Intuitively, advanced composition "rescales" how privacy accumulates: instead of ε growing in proportion to k, it grows on the order of $\sqrt{k}$ for large k, and for a given target overall δ (Dwork et al., 2010).

In simpler terms, you can apply more mechanisms for the same privacy budget by accepting a tiny increase in the failure probability parameter. The formal theorem (Dwork et al., 2010) shows that for (ε, δ)-DP mechanisms repeated k times, the total privacy can be bounded by approximately:

$$\left(\varepsilon \sqrt{2k * ln\left(\frac{1}{\delta}\right)} + k\varepsilon^2, k\delta + \delta' \right) \text{ for any } \delta' > 0$$

Relevant info is that ε grows much more slowly than k which makes the guarantee significantly tighter than in basic composition.

7.2 Rényi differential privacy

Beyond (ε, δ) accounting, more advanced methods have been developed that apply alternative measures of privacy loss. Rényi Differential Privacy (RDP) is a framework introduced by (Mironov, 2017) that generalizes DP using Rényi divergence (van Erven et al., 2012). RDP is parameterized by an order α and corresponds to measuring the moments of the privacy loss distribution. It provides a unified view encompassing methods from pure to approximate DP.

Rényi Divergence: The Rényi divergence $D_\alpha(P||Q)$ of order α between probability distributions P and Q is:

$$D_\alpha(P||Q) = \frac{1}{\alpha - 1} ln\mathbb{E}_{x\sim Q}\left[\left(\frac{P(x)}{Q(x)}\right)^\alpha\right]$$

Where

$\alpha \geq 1$ is the order parameter. When $\alpha \to \infty$, RDP approach pure-DP

RDP's key benefit is composition: RDP parameters add directly. If a mechanism has $(\alpha, \varepsilon_\alpha)$ −RDP (meaning the α-th moment of privacy loss is bounded by ε_α) then k independent uses have at most $(\alpha, k\varepsilon_\alpha)$-RDP. This additive property simplifies privacy loss accounting over multiple iterations. After composition, the RDP guarantee can be converted back to (ε, δ)-DP with:

$$\varepsilon = \varepsilon_\alpha + \frac{\ln\left(\frac{1}{\delta}\right)}{\alpha - 1}$$

Which has been improved with tighter bounds (Balle et al., 2018) as:

$$\varepsilon = \varepsilon_1 + \frac{\ln(\delta) + \ln\left(1 - \frac{1}{\delta}\right) - \ln(\alpha)}{\alpha - 1}$$

RDP accounting provides tighter bounds than basic or advanced composition by using the complete privacy loss distribution rather than worst-case bounds at each step. It's particularly useful for mechanism sequences like in ML training, enabling privacy loss tracking across all possible δ values simultaneously.

7.3 Moments accountant and analytical moments accountant

Moments accountant is a technique introduced by Abadi et al. (2016) alongside the DP-SGD algorithm. It is closely related to RDP – in fact, it was a precursor that inspired RDP's development**.**

The moments accountant keeps track of the log moments of the PLRV at each step, and thus can tightly bound the overall (ε, δ) after many compositions. This method was tailored for the subsampled Gaussian mechanism this work utilises by bounding higher-order moments and it achieves much tighter accounting than even the advanced composition theorem. As a formulation, Moments Accountant can be represented as optimized minimization task:

$$\varepsilon = \min_{\alpha}\left(\varepsilon_{\alpha} + \frac{ln\left(\frac{1}{\delta}\right)}{\alpha - 1}\right)$$

Where α represents different moment orders.

However, it lacks an explicit, standalone definition of privacy and is primarily a practical tool for moment-based privacy accounting rather than a formal privacy definition.

While the original Moments Accountant often required numerical approximations to compute log moments and bound the cumulative privacy loss, subsequent research introduced closed-form analytical expressions known as the Analytical Moments Accountant (AMA) (Y. X. Wang et al., 2018). Specifically, AMA utilizes exact formulas derived from RDP theory for certain key mechanisms—most notably the subsampled Gaussian mechanism commonly used in DP-SGD. Formally, AMA directly computes the Rényi divergence of order α for the subsampled Gaussian mechanism. Given sampling probability q and noise scale σ, closed-form expression for the RDP bound can be derived as follows (Y. X. Wang et al., 2018):

$$\varepsilon_{\alpha}(q,\sigma) \leq \frac{1}{\alpha - 1}\ln\left(1 + q^{2}\binom{\alpha}{2}\min\left\{4\left(e^{\frac{1}{\sigma^{2}}} - 1\right), 2e^{\frac{1}{\sigma^{2}}}\right\} + \sum_{j=3}^{\alpha} q^{j}\binom{\alpha}{j} 2^{j} e^{\frac{j}{\sigma^{2}}}\right)$$

These this exact bound AMA accumulates RDP parameters over k steps through simple addition, given that RDP composes linearly:

$$\varepsilon_{\alpha}^{(total)}(k,q,\sigma) = k \times \varepsilon_{\alpha}(q,\sigma)$$

from which the precise (ε, δ)-DP guarantee can be obtained by minimizing over different orders after k steps as seen with MA:

$$\varepsilon = \min_{\alpha}\left(\varepsilon_{\alpha}^{(total)}(k,q,\sigma) + \frac{ln\left(\frac{1}{\delta}\right)}{\alpha - 1}\right)$$

By explicitly computing privacy loss through these closed-form solutions, AMA significantly reduces pessimism inherent in earlier numerical approximations. This yields tighter privacy accounting—providing lower and more realistic estimates of the privacy budget compared to previous methods.

8 Experiment setup and results

Experiments were conducted using modified FeTS Challenge codebase (FeTS Challenge) where necessary modifications were made to the codebase to provide the desired simulation pipeline and settings.

8.1 Simulation settings

Experiments were run as a single GPU (Tesla V100 32GB) simulation in DGX-1 computing unit (NVIDIA DGX-1). This multi-GPU setup provided the possibility of running multiple independent simulations in parallel. Institution split was set as “Partitioning 2” (Figure 16) which initializes 33 clients with local training dataset size N ranging from 3 to 136 patients.

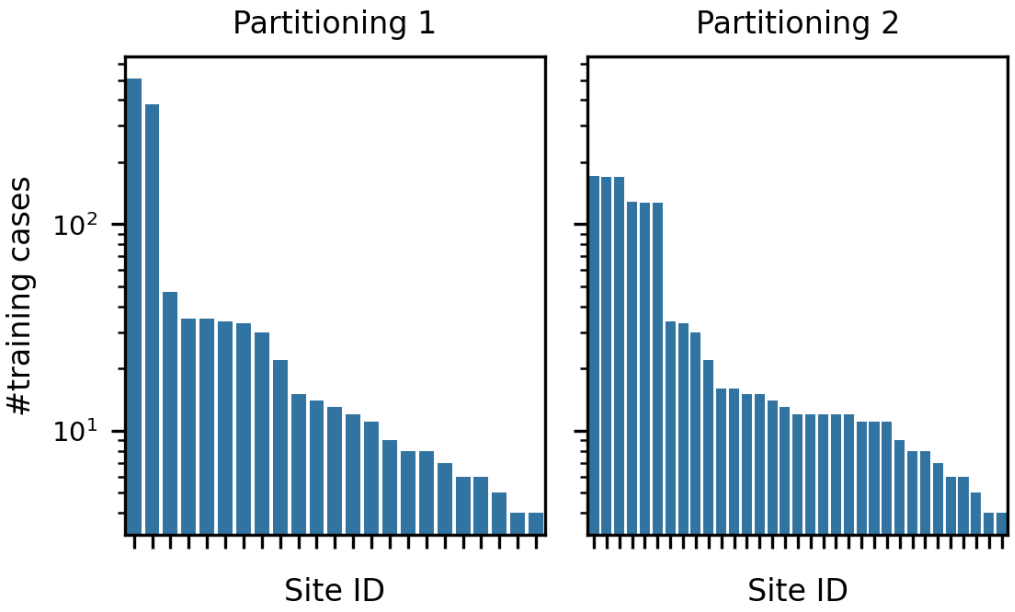

Figure 16. Client partitioning and patient counts (FeTS 2024 Challenge).

Batch size of $b = 1$ patients was used for all subsampling methods and sampling ratio for each client was set at $q = \frac{b}{N}$. Local training epoch follows roughly DP-SGD with per-patient sampling (Figure 17):

Algorithm 1 Local Training with Subsampled Gaussian Differential Privacy (Client-side)

Require:

Global model parameters θ_G,
Local epochs E,
local dataset $\mathcal{D}$,
sampling ratio q,
gradient clip bound C,
noise scale σ,
patches per volume P
function SUBSAMPLEDGAUSSIAN(θ_G, E, $\mathcal{D}$, q, C, σ, P)
    Initialize local model parameters $\theta_L \leftarrow \theta_G$
    Initialize accumulated gradients $\mathcal{G} \leftarrow \{\}$
    for each local epoch E **do**
        $\mathcal{D}_s \leftarrow \text{subSample}(\mathcal{D}, q)$
        for each patient $p \in \mathcal{D}_s$ **do**
            for $i = 1$ to P **do**
                Calculate loss $\mathcal{L}(\theta_L, p_i)$
                Compute gradient $g \leftarrow \nabla_{\theta_L}\mathcal{L}(\theta_L, p_i)$
                Clip gradient $\hat{g} \leftarrow g \div \max\left(1, \frac{\|g\|_2}{C}\right)$
                $\mathcal{G} \leftarrow \text{accumulateGrads}(\mathcal{G}, \hat{g})$
            end for
        end for
        for each parameter $\phi \in \mathcal{G}$ **do**
            Generate noise $\eta \sim \mathcal{N}(0, \sigma^2 C^2 \mathbf{I})$
            $\phi \leftarrow \frac{\phi+\eta}{P}$ Average over P patches
        end for
        Update local model θ_L using noisy accumulated gradients $\mathcal{G}$
    end for
    return θ_L

Figure 17. Local training with subsampled Gaussian differential privacy.

Each client initiates training using global model parameters Θ_G and then proceeds to locally train these parameters on their own dataset $\mathcal{D}$. During each local epoch E, the dataset undergoes subsampling at a defined ratio q, after which training is performed for each patient individually. For every patient, we extract $P = 40$ patches with all four sequences resulting per-patch input shape of [1,4,64,64,64]. Per-patch gradients g are computed based on the model's loss. These per-patch gradients are subsequently clipped to a predefined norm bound C, effectively controlling sensitivity and ensuring stable gradient magnitudes. Clipped gradients are accumulated throughout the patient's

training steps. Before updating the local model, Gaussian noise with scale σC is sampled and added to accumulated gradients to enforce differential privacy. Finally, the noisy accumulated gradients are averaged over the number of patches P, before applying them to the local model parameters.

σ, C, δ, learning rate (lr) and local epochs are pre-set for all clients for each simulation (Table 1):

Privacy Setting	FL Rounds	Client selection	C	σ	δ	ε	lr	Local epochs
Baseline	30	20%	–	–	–	–	$5E-5$	1
low	30	20%	0.5	0.5	$1E-5$	4.84	$5E-3$	5
Medium	30	20%	0.5	1.0	$1E-5$	2.42	$5E-3$	5
High	30	20%	0.5	2.0	$1E-5$	1.21	$5E-3$	5
High+	30	20%	0.5	2.5	$1E-5$	0.96	$5E-3$	5

Table 1. Simulation settings.

Baseline represents the default non-DP simulation where model is updated after each example and each local training participation consists of 1 epoch. Since our DP method applies fewer model updates per epoch, we have increased local epoch from 1 to 5 for DP-simulations.

Usually, σ can be calculated dynamically based on the given ε and current remaining budged. For simplicity and testing consistency, we set σ at fixed value used in every iteration of the DP mechanism and calculate total ε with different accounting methods after the simulation is complete.

Each simulation with their specific settings were run for 30 rounds and client selection was set to include 20% of the total clients to train per round.

Each local DP-training round uses fixed $lr = 5E-3$ which is increased from baseline to account for the fewer but larger (per-example vs per-user) model updates. Privacy Accounting results were calculated after simulations were finished with the help of tracked interation count T. Initial ε is the budget for

single iteration calculated from the initial simulation settings e.g. if we would have set fixed privacy budget instead of noise scale, this value would correspond to these settings.

8.2 Results

The experimental results provide insights into the trade-off between segmentation performance and patient-level local differential privacy. The performance of the 3D Residual U-Net model under different privacy settings ('low' $\sigma = 0.5$, 'medium' $\sigma = 1.0$, 'high' $\sigma = 2.0$, 'high+' $\sigma = 2.5$) was evaluated against a non-private baseline using Dice Similarity Coefficient (DSC) and 95th percentile Hausdorff Distance (H95) metrics.

8.2.1 Overall performance

Figure 18 illustrates the convergence of the mean Dice score across clients over the 30 federated learning rounds for each simulation setting. The baseline (non-DP) model achieves the highest final Dice score (0.7738 at round 24), establishing the benchmark performance. Introducing LDP results in a decrease in the overall Dice score, as expected.

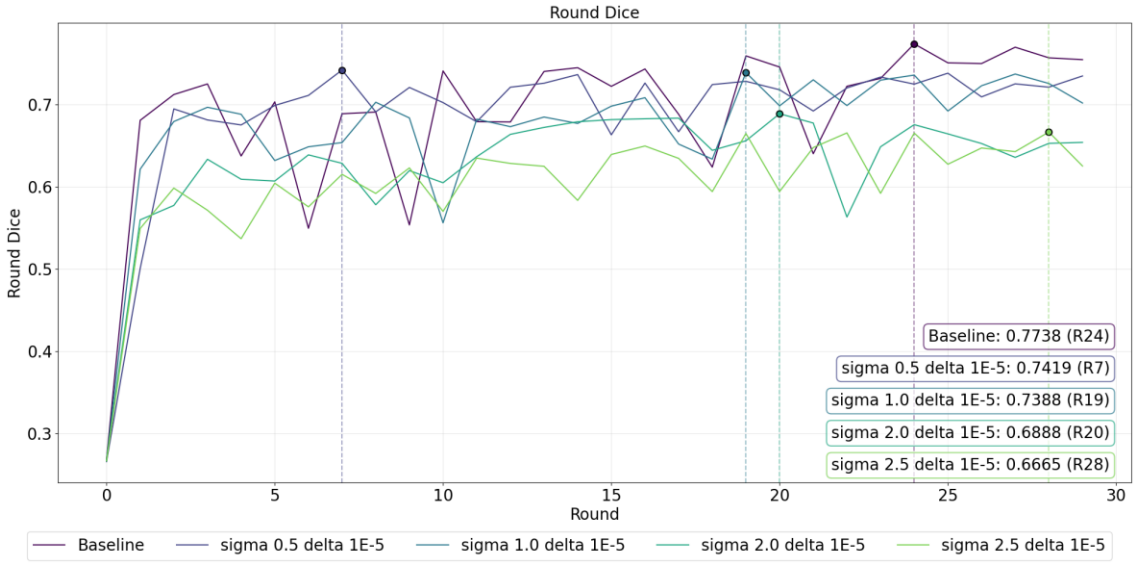

Figure 18. Round Dice per simulation and best score.

The magnitude of this performance reduction correlates strongly with the privacy level: higher noise scales (σ), corresponding to stricter privacy guarantees (lower ε), lead to lower final Dice scores. The 'low' privacy setting ($\sigma = 0.5$) shows the least impact, achieving a final Dice score close to the baseline, while the 'High+' setting ($\sigma = 2.5$) exhibits the most pronounced drop. Despite the inherent noise, all LDP models demonstrate stable convergence during training, albeit potentially reaching lower final performance levels compared to the non-private baseline.

8.2.2 Segmentation accuracy per tumor sub-region.

Figures 19 and 20 provide a more granular view of the performance impact on specific tumor sub-regions: Healthy tissue (label 0), ET (label 1), TC (label 2) WT (label 4).

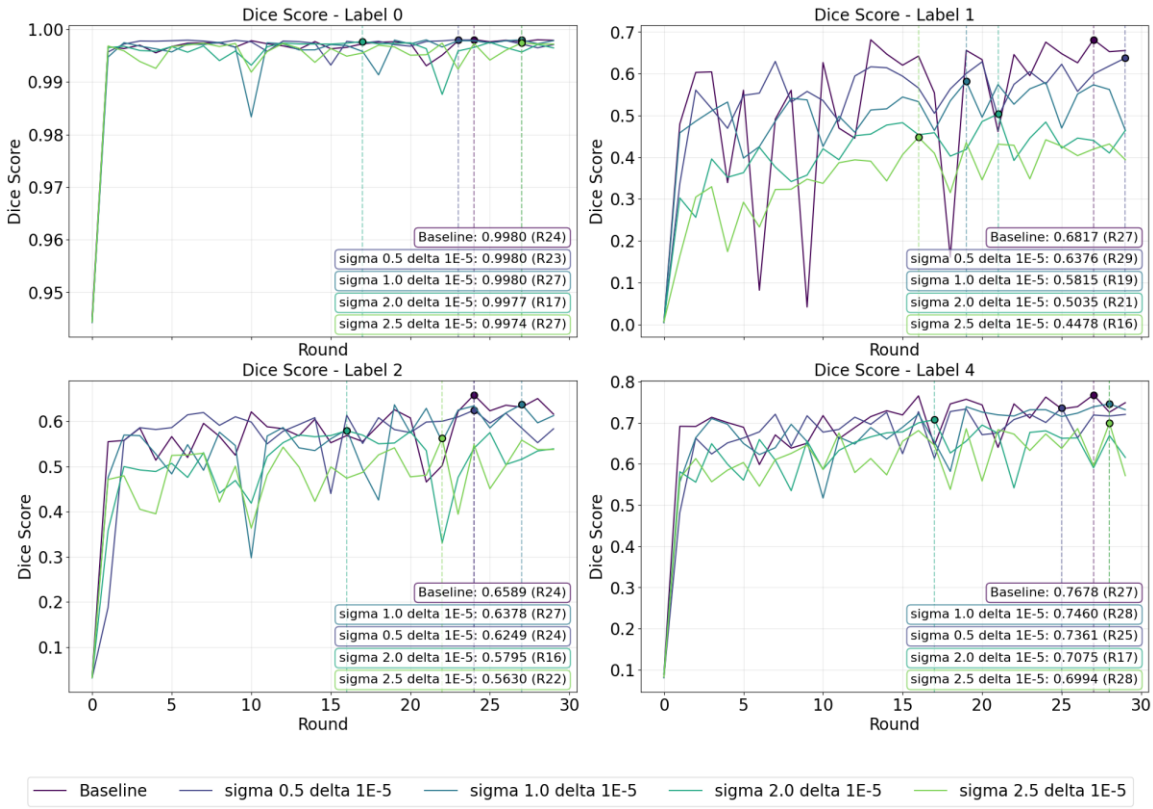

Figure 19. Dice Score per predicted class, higher is better.

For Dice scores (Figure 19), the baseline model consistently achieves the highest Dice scores across all sub-regions. Introducing LDP reduces the Dice score for all regions, with the reduction generally increasing as the privacy level increases (from 'low' to 'High+'). The impact appears most pronounced for the Enhancing Tumor (ET), which is most tricky region to predict accurately from area and shape point of view.

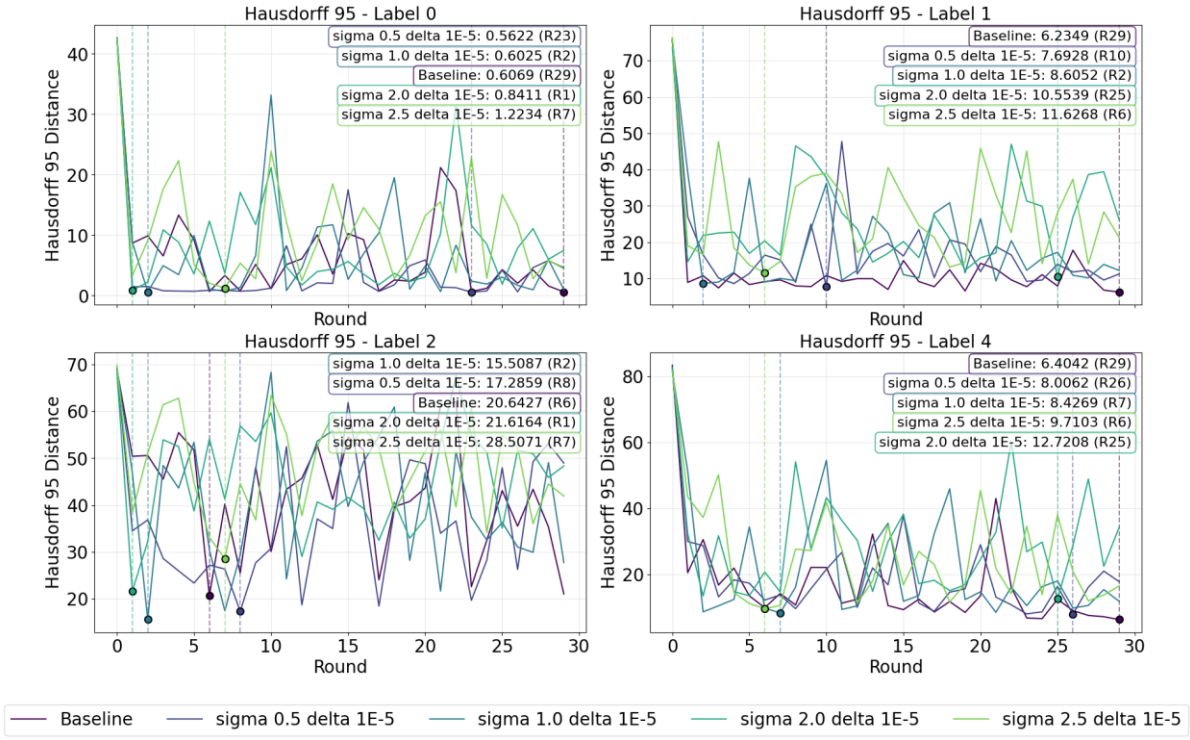

Figure 20. H95 score per predicted label, lower is better.

The H95 (Figure 20), sensitive to boundary delineation errors, shows a similar trend. Lower H95 scores indicate better boundary agreement (lower is better). The baseline achieves mostly the lowest H95 scores. Increasing the noise scale (σ) generally leads to higher H95 scores (worse boundary accuracy) across ET, TC, and WT, indicating that the noise affects the model's ability to precisely capture tumor boundaries. Again, the 'High+' setting shows the largest increase in H95 distance compared to the baseline. We can also observe that the randomness introduced by our client selection method affects round by round performances from the large variations in client dataset sizes. This results some cases, where DP simulations outperform baseline in some labels with low privacy regimes ($\sigma = 0.5, 1.0$), especially in H95 metrics where the baseline scores are really close to 0.

Figure 21 provides a clearer view of the relative performance change by showing the percentage difference from the baseline's best scores for both Dice and H95 metrics.

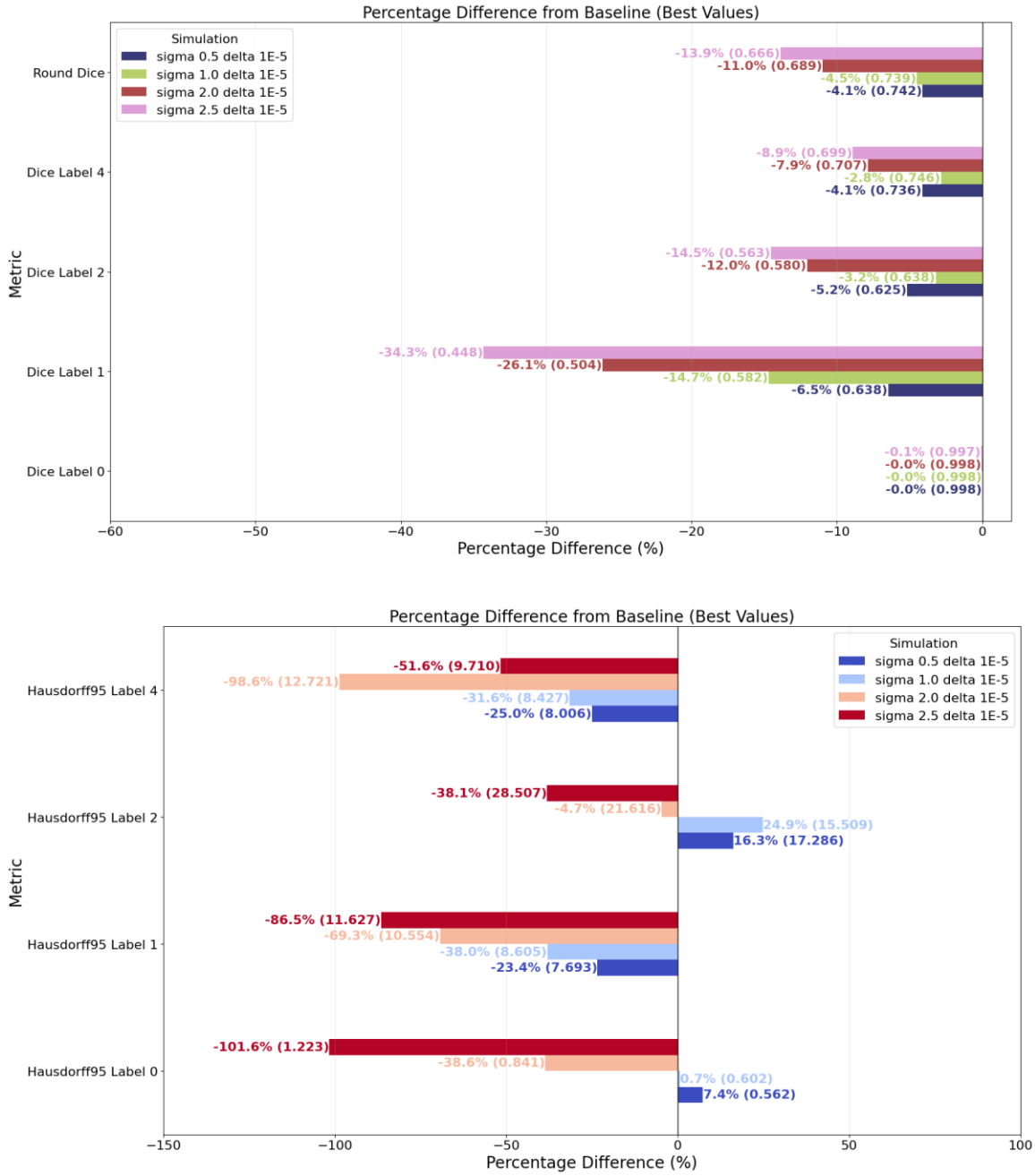

Figure 21. Dice and H95 scores, % from baseline.

The Dice score for healthy tissue (Label 0) remains largely unaffected, staying within 1% of the baseline across all privacy settings. For tumor regions, the 'low' privacy setting ($\sigma = 0.5$) results in Dice score reductions ranging from approximately 4.1% (Label 4) to 6.5% (Label 1) compared to the baseline.

Increasing privacy strengthens this trend. In the 'High+' setting ($\sigma = 2.5$), the Dice score drops by roughly 8.9% for label 4, 14.5% for label 2, and a more substantial 34.3% for label 1 relative to the baseline. This highlights that the more challenging regions suffers the most significant relative utility loss under strong LDP.

The H95 metric (Figure 21), sensitive to boundary errors, shows a more dramatic impact from LDP on paper. Even the 'low' privacy setting leads to significant percentage changes in H95 compared to the baseline, indicating a notable degradation in boundary accuracy. In reality, since H95 gets close to 0 in some labels, even the $\sim 100\%$ drop in performance is still barely over 1.0 in actual score and can be argued to be non-issue. This further suggests that while boundary delineation is highly sensitive to initial noise introduction, further increases in noise scale have a diminishing relative impact on the H95 scores.

8.2.3 Privacy accounting.

Privacy accounting analyses were conducted to evaluate the cumulative privacy guarantees achieved throughout the iterative federated training process under various LDP settings. Figures 22 to 25 illustrate the privacy budgets calculated across different clients using previously mentioned accounting methodologies: Basic Composition (BC), Advanced Composition (AC), Rényi Differential Privacy (RDP), Moments Accountant (MA) and Analytical Moments Accountant (AMA).

Figure 22 reveals how client dataset sizes influence both iteration counts and sampling rates. While larger datasets naturally resulted in more iterations per client, the sampling rate was also samallest in clients with largest datasets.

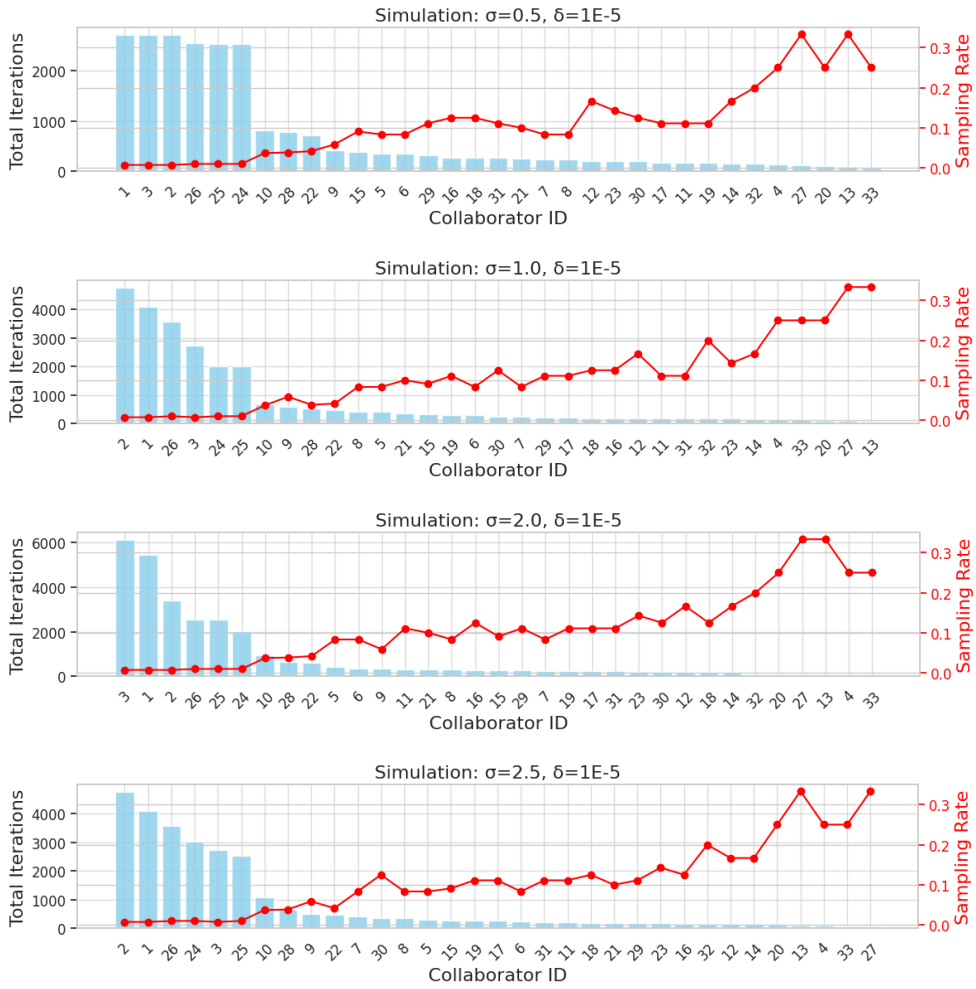

Figure 22. Iteration count and sampling rate of each client.

Although, intuitively larger iteration counts could suggest higher overall privacy budgets, the low sampling rate effectively counterbalanced this, demonstrating efficient budget management even in larger datasets.

For analyzing total spent privacy budgets per client, Figures 23 and 24 clearly show that advanced accounting methods yield significantly tighter privacy bounds compared to basic composition.

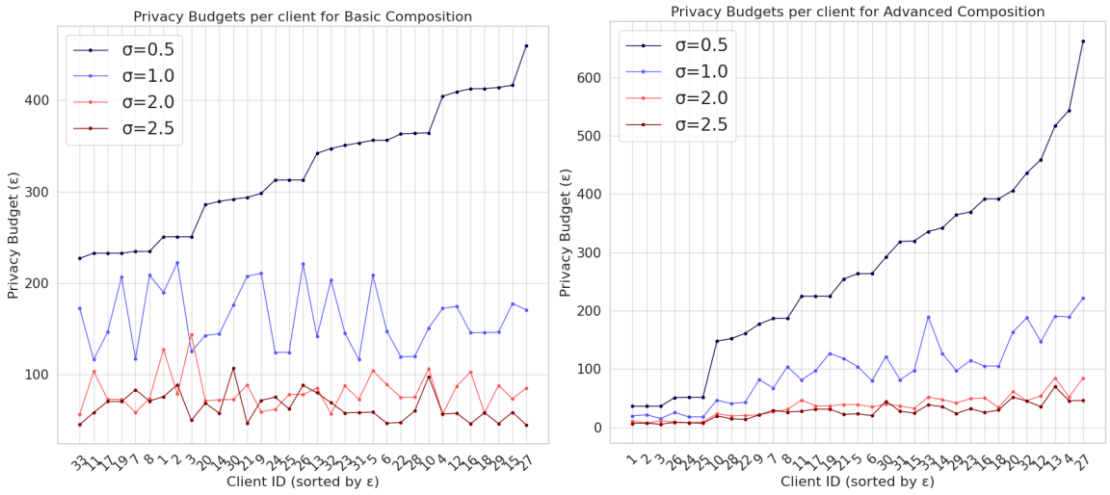

Figure 23. Privacy budgets per client with Basic Composition and Advanced Composition.

Under basic composition, even the 'high' privacy settings showed substantially inflated privacy budgets. Total budgets became increasingly pronounced at low noise scales where total budget was in the hundreds across all clients. In contrast, advanced composition provided notably improved bounds by incorporating sublinear growth of cumulative privacy loss.

More refined privacy accounting techniques, including Rényi Differential Privacy (RDP) and Moments Accountant (MA) (Figure 24), offered further tightening of privacy guarantees.

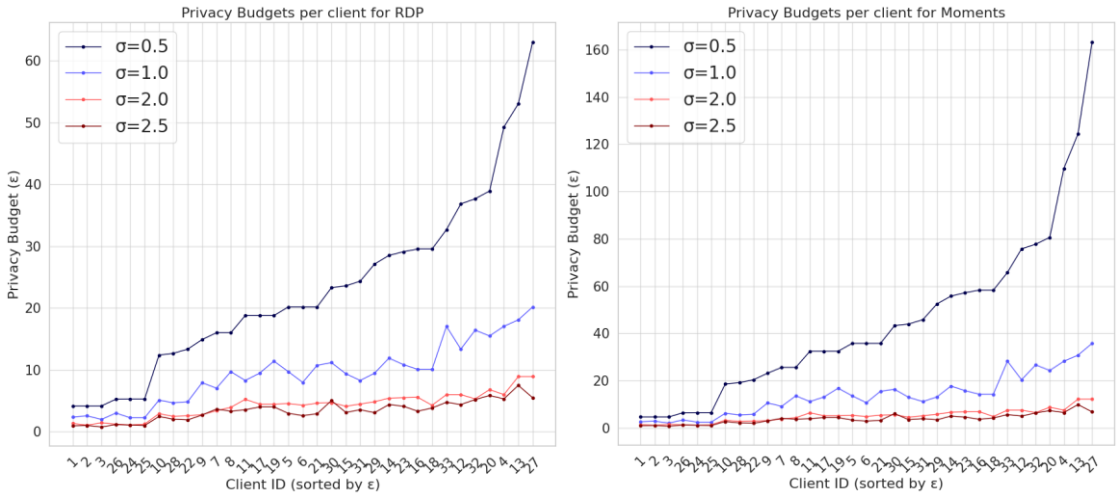

Figure 24. Privacy budgets per client with RDP and Moments Accountant.

Moments Accountant, leveraging the distributional characteristics of privacy loss, consistently delivered lower cumulative privacy estimates compared to basic and advanced composition or even RDP.

The Analytical Moments Accountant (AMA), depicted in Figure 25, presented the tightest overall privacy guarantees across all privacy settings.

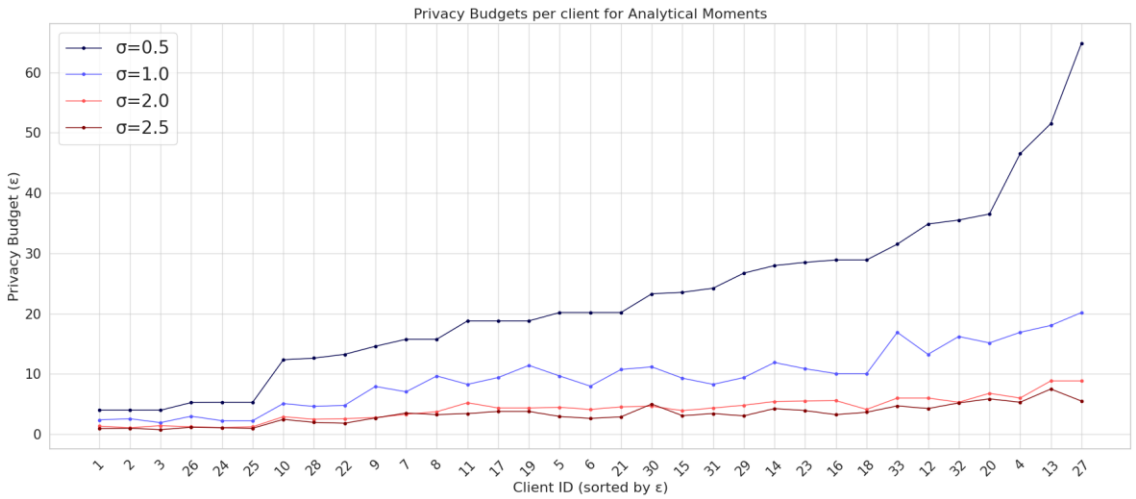

Figure 25. Privacy budget per client with Analytical Moments.

AMA provided extremely accurate bounds for subsampled Gaussian mechanisms across all clients, highlighting the benefit of precise, closed-form computations over numerical approximations.

Table 2 summarizes the best client-level ε values achieved under each privacy setting, clearly illustrating the effective management of privacy budgets even at lower privacy levels. Specifically, despite some clients having small datasets and reduced privacy amplification via subsampling, MA and AMA accounting maintained ε within acceptable privacy thresholds, indicating practical feasibility for rigorous patient-level privacy protection in federated medical imaging contexts even with lower noise scales.

Client ID	sigma 0.5 ε	sigma 1.0 ε	sigma 2.0 ε	sigma 2.5 ε	Global Min ε
3	3.9744	1.9137	1.4217	0.7485	0.7485
1	3.9744	2.364	1.3381	0.9193	0.9193
25	5.2703	2.2343	1.2247	0.9748	0.9748
2	3.9744	2.563	1.0518	0.9944	0.9944
24	5.2703	2.2343	1.0926	1.0698	1.0698
26	5.2424	2.9825	1.2185	1.1521	1.1521
22	13.2419	4.7971	2.5791	1.8189	1.8189
28	12.6169	4.5933	2.4729	1.9591	1.9591
10	12.3391	5.0804	2.8938	2.4549	2.4549
6	20.1753	7.9695	4.1063	2.6113	2.6113
9	14.58	7.9174	2.7496	2.691	2.691
21	20.1753	10.7302	4.5308	2.8745	2.8745
5	20.1753	9.6553	4.4631	2.9372	2.9372
29	26.7103	9.3932	4.7971	3.0391	3.0391
15	23.5285	9.2922	3.9017	3.0755	3.0755
16	28.8978	10.0459	5.5662	3.2351	3.2351
8	15.7455	9.6553	3.7238	3.2351	3.2351
7	15.7455	7.0284	3.3059	3.5114	3.3059
31	24.2103	8.2647	4.346	3.4213	3.4213
11	18.7864	8.2647	5.2189	3.4213	3.4213
18	28.8978	10.0459	4.1063	3.6431	3.6431
19	18.7864	11.4107	4.346	3.7712	3.7712
17	18.7864	9.3932	4.346	3.7712	3.7712
23	28.4961	10.8552	5.5041	3.9135	3.9135
14	27.9603	11.8877	5.4186	4.2522	4.2522
12	34.8463	13.2419	5.9902	4.2522	4.2522
30	23.2728	11.1677	4.6323	5.0026	4.6323
33	31.5129	16.8913	5.9902	4.6929	4.6929
32	35.5129	16.2038	5.2988	5.1826	5.1826
4	46.5129	16.8913	5.9902	5.2988	5.2988
27	63.0259	20.1753	8.8376	5.4904	5.4904
20	36.5129	15.1315	6.7811	5.8584	5.8584
13	51.5129	18.0371	8.8376	7.4894	7.4894

Table 2. Best client ε per simulation (lower is better).

These results collectively demonstrate that having large datasets define great base for robust DP implementations. In addition, careful selection and optimization of privacy parameters, supported by advanced privacy accounting methods, enable strong privacy guarantees with minimal detriment to segmentation performance. Such precise accounting is crucial for maintaining regulatory compliance and ethical standards in collaborative healthcare research.

9 Conclusion

The objective of this thesis was to investigate the feasibility of implementing local differential privacy within federated learning for medical image segmentation, specifically brain tumor segmentation, addressing critical issues surrounding patient data privacy. This study successfully demonstrated that local differential privacy, applied through Gaussian noise addition and gradient clipping on client-side updates, is a viable approach to enhance patient-level confidentiality in federated machine learning models.

Our experiments revealed that integrating local differential privacy slightly compromises segmentation performance, reflected by marginally decreased accuracy metrics such as the Dice similarity coefficient and Hausdorff distance. However, careful optimization of privacy parameters—namely, σ, C, q and local training epochs—significantly mitigated the performance decline. The trade-off between privacy and accuracy was effectively balanced by employing rigorous privacy accounting methods, notably the Moments Accountant and Analytical Moments Accountant, which precisely tracked the cumulative privacy budgets across iterative federated training rounds.

These findings illustrate that despite a modest computational overhead and complexity in hyperparameter tuning, LDP can be effectively integrated into cross-silo federated learning setups common in healthcare. The approach protects patient-sensitive data robustly against inference attacks, thus aligning with stringent regulatory frameworks governing medical data privacy.

Future work should explore further optimizations in privacy mechanisms, such as adaptive privacy budgets, gradient clipping- and alternative noise-strategies, to further improve performance without compromising privacy guarantees. Extending this approach to larger datasets and diverse medical imaging tasks could solidify its practical viability. Ultimately, this thesis confirms that local differential privacy can play a critical role in securely advancing federated learning applications in medical imaging, paving the way for safer and broader adoption of collaborative AI solutions in healthcare.

References

Abadi, M., McMahan, H. B., Chu, A., Mironov, I., Zhang, L., Goodfellow, I., & Talwar, K. (2016). Deep Learning with Differential Privacy. *Proceedings of the ACM Conference on Computer and Communications Security*, *24-28-October-2016*, 308–318. doi: 10.1145/2976749.2978318

Abowd, J. M., & Hawes, M. B. (2022). *Confidentiality Protection in the 2020 US Census of Population and Housing*. doi: 10.1146/annurev-statistics-010422-034226

Antunes, R. S., Da Costa, C. A., Küderle, A., Yari, I. A., & Eskofier, B. (2022). Federated Learning for Healthcare: Systematic Review and Architecture Proposal. *ACM Transactions on Intelligent Systems and Technology (TIST)*, *13*(4). doi: 10.1145/3501813

Balle, B., Barthe, G., & Gaboardi, M. (2018). Privacy Amplification by Subsampling: Tight Analyses via Couplings and Divergences. *Advances in Neural Information Processing Systems*, *2018-December*, 6277–6287. Retrieved from https://arxiv.org/abs/1807.01647v2

Bitar, R., Leung, G., Perng, R., Tadros, S., Moody, A. R., Sarrazin, J., McGregor, C., Christakis, M., Symons, S., Nelson, A., & Roberts, T. P. (2006). MR pulse sequences: what every radiologist wants to know but is afraid to ask. *Radiographics : A Review Publication of the Radiological Society of North America, Inc*, *26*(2), 513–537. doi: 10.1148/RG.262055063

Brendan McMahan, H., Moore, E., Ramage, D., Hampson, S., & Agüera y Arcas, B. (2016). Communication-Efficient Learning of Deep Networks from Decentralized Data. *Proceedings of the 20th International Conference on Artificial Intelligence and Statistics, AISTATS 2017*. Retrieved from https://arxiv.org/abs/1602.05629v4

Dalenius, T. (1977). *Towards a methodology for statistical disclosure control*. Statistics Sweden. Retrieved from https://hdl.handle.net/1813/111303

Despotović, I., Goossens, B., & Philips, W. (2015). MRI Segmentation of the Human Brain: Challenges, Methods, and Applications. *Computational*

and Mathematical Methods in Medicine, *2015*, 450341. doi: 10.1155/2015/450341

D'este, S. H., Nielsen, M. B., & Hansen, A. E. (2021). Visualizing Glioma Infiltration by the Combination of Multimodality Imaging and Artificial Intelligence, a Systematic Review of the Literature. *Diagnostics*, *11*(4), 592. doi: 10.3390/DIAGNOSTICS11040592

Dwork, C. (2006). Differential Privacy. *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, *4052 LNCS*, 1–12. doi: 10.1007/11787006_1

Dwork, C., & Roth, A. (2014). The Algorithmic Foundations of Differential Privacy. *Foundations and Trends® in Theoretical Computer Science*, *9*(3–4), 211–407. doi: 10.1561/0400000042

Dwork, C., Rothblum, G. N., & Vadhan, S. (2010). Boosting and differential privacy. *Proceedings - Annual IEEE Symposium on Foundations of Computer Science, FOCS*, 51–60. doi: 10.1109/FOCS.2010.12

Essential Instrument for AI Research | NVIDIA DGX-1. (n.d.). Retrieved from https://www.nvidia.com/en-in/data-center/dgx-1/

FeTS 2024 Challenge. (n.d.). Retrieved from https://www.synapse.org/Synapse:syn54079892/wiki/626854

GitHub - FeTS-AI/Challenge: The repo for the FeTS Challenge. (n.d.). Retrieved from https://github.com/FeTS-AI/Challenge

Hao, X., Xu, D., Bansal, R., Dong, Z., Liu, J., Wang, Z., Kangarlu, A., Liu, F., Duan, Y., Shova, S., Gerber, A. J., & Peterson, B. S. (2011). Multimodal magnetic resonance imaging: The coordinated use of multiple, mutually informative probes to understand brain structure and function. *Human Brain Mapping*, *34*(2), 253. doi: 10.1002/HBM.21440

Hausdorff, F. (1914). *Grundzüge der mengenlehre* (Vol. 7). von Veit.

He, K., Zhang, X., Ren, S., & Sun, J. (2015). Deep Residual Learning for Image Recognition. *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, *2016-December*, 770–778. doi: 10.1109/CVPR.2016.90

Hesamian, M. H., Jia, W., He, X., & Kennedy, P. (2019). Deep Learning Techniques for Medical Image Segmentation: Achievements and Challenges. *Journal of Digital Imaging*, *32*(4), 582. doi: 10.1007/S10278-019-00227-X

Huang, C., Huang, J., & Liu, X. (2022). *Cross-Silo Federated Learning: Challenges and Opportunities*. Retrieved from https://arxiv.org/abs/2206.12949v1

Illich, I. (1971). *Deschooling Society*. New York: Harper & Row.

Isensee, F., Jaeger, P. F., Kohl, S. A. A., Petersen, J., & Maier-Hein, K. H. (2021). nnU-Net: a self-configuring method for deep learning-based biomedical image segmentation. *Nature Methods*, *18*(2), 203–211. doi: 10.1038/S41592-020-01008-Z

Kairouz, P., McMahan, H. B., Avent, B., Bellet, A., Bennis, M., Bhagoji, A. N., & Zhao, S. (2021). Advances and Open Problems in Federated Learning. *Foundations and Trends in Machine Learning*, *14*(1–2), 1–210.

Khan, M. I., Azeem, M. A., Alhoniemi, E., Kontio, E., Khan, S. A., & Jafaritadi, M. (2023). *Regularized Weight Aggregation in Networked Federated Learning for Glioblastoma Segmentation*. Retrieved from https://arxiv.org/abs/2301.12617v1

Khan, M. I., Kontio, E., Khan, S. A., & Jafaritadi, M. (2024). *Recommender Engine Driven Client Selection in Federated Brain Tumor Segmentation*. Retrieved from https://arxiv.org/abs/2412.20250v1

Learning with Privacy at Scale Differential Privacy Team, Apple. (n.d.). Retrieved from https://docs-assets.developer.apple.com/ml-research/papers/learning-with-privacy-at-scale.pdf

McDonald, R. J., Schwartz, K. M., Eckel, L. J., Diehn, F. E., Hunt, C. H., Bartholmai, B. J., Erickson, B. J., & Kallmes, D. F. (2015). The effects of changes in utilization and technological advancements of cross-sectional imaging on radiologist workload. *Academic Radiology*, *22*(9), 1191–1198. doi: 10.1016/J.ACRA.2015.05.007

Menze, B. H., Jakab, A., Bauer, S., Kalpathy-Cramer, J., Farahani, K., Kirby, J., Burren, Y., Porz, N., Slotboom, J., Wiest, R., Lanczi, L., Gerstner,

E., Weber, M. A., Arbel, T., Avants, B. B., Ayache, N., Buendia, P., Collins, D. L., Cordier, N., … Van Leemput, K. (2015). The Multimodal Brain Tumor Image Segmentation Benchmark (BRATS). *IEEE Transactions on Medical Imaging*, *34*(10), 1993–2024. doi: 10.1109/TMI.2014.2377694

Mironov, I. (2017). Renyi Differential Privacy. *Proceedings - IEEE Computer Security Foundations Symposium*, 263–275. doi: 10.1109/CSF.2017.11

Molinuevo, D., Fóti, K., & Kruse, F. (2017). *Delivering hospital services a greater role for the private sector?* 61. Retrieved from https://www.eurofound.europa.eu/en/publications/2017/delivering-hospital-services-greater-role-private-sector#tab-01

Nasr, M., Shokri, R., & Houmansadr, A. (2019). Comprehensive privacy analysis of deep learning: Passive and active white-box inference attacks against centralized and federated learning. *Proceedings - IEEE Symposium on Security and Privacy*, *2019-May*, 739–753. doi: 10.1109/SP.2019.00065

Pascanu, R., Mikolov, T., & Bengio, Y. (2012). On the difficulty of training Recurrent Neural Networks. *30th International Conference on Machine Learning, ICML 2013*, *PART 3*, 2347–2355. Retrieved from https://arxiv.org/abs/1211.5063v2

Pati, S., Baid, U., Zenk, M., Edwards, B., Sheller, M., Reina, G. A., Foley, P., Gruzdev, A., Martin, J., Albarqouni, S., Chen, Y., Shinohara, R. T., Reinke, A., Zimmerer, D., Freymann, J. B., Kirby, J. S., Davatzikos, C., Colen, R. R., Kotrotsou, A., … Bakas, S. (2021a). *The Federated Tumor Segmentation (FeTS) Challenge*. Retrieved from https://arxiv.org/abs/2105.05874v2

Pati, S., Baid, U., Zenk, M., Edwards, B., Sheller, M., Reina, G. A., Foley, P., Gruzdev, A., Martin, J., Albarqouni, S., Chen, Y., Shinohara, R. T., Reinke, A., Zimmerer, D., Freymann, J. B., Kirby, J. S., Davatzikos, C., Colen, R. R., Kotrotsou, A., … Bakas, S. (2021b). *The Federated Tumor Segmentation (FeTS) Challenge*. Retrieved from https://arxiv.org/abs/2105.05874v2

Ronneberger, O., Fischer, P., & Brox, T. (2015). 2015-U-Net. *ArXiv*, 1–8. Retrieved from http://lmb.informatik.uni-freiburg.de/%0Aarxiv:1505.04597v1

Shan, F., Mao, S., Lu, Y., & Li, S. (n.d.). Differential Privacy Federated Learning: A Comprehensive Review. *IJACSA) International Journal of Advanced Computer Science and Applications*, *15*(7), 2024. Retrieved from www.ijacsa.thesai.org

Sørensen, T. (1948). A method of establishing groups of equal amplitude in plant sociology based on similarity of species and its application to analyses of the vegetation on Danish commons. *Biol Skrifter/Kongelige Danske Videnskabernes Selskab.*, *5*, 1.

Teo, Z. L., Jin, L., Li, S., Miao, D., Zhang, X., Ng, W. Y., Tan, T. F., Lee, D. M., Chua, K. J., Heng, J., Liu, Y., Goh, R. S. M., & Ting, D. S. W. (2024). Federated machine learning in healthcare: A systematic review on clinical applications and technical architecture. *Cell Reports. Medicine*, *5*(2). doi: 10.1016/J.XCRM.2024.101419

toisiolaki 552/2019. (2019, April 19). *laki sosiaali- ja terveystietojen toissijaisesta käytöstä 552/2019*. Retrieved from https://www.finlex.fi/fi/laki/alkup/2019/20190552#

Tramèr, F., Zhang, F., Juels, A., Reiter, M. K., & Ristenpart, T. (2016). Stealing Machine Learning Models via Prediction APIs. *Proceedings of the 25th USENIX Security Symposium*, 601–618. Retrieved from https://arxiv.org/abs/1609.02943v2

U-Net: A Comprehensive Guide to Its Architecture and Applications - viso.ai. (n.d.). Retrieved from https://viso.ai/deep-learning/u-net-a-comprehensive-guide-to-its-architecture-and-applications/#elementor-toc__heading-anchor-27

Union, P. O. of the E. (2025). *Regulation (EU) 2017/745 of the European Parliament and of the Council of April 2017 on medical devices, amending Directive 2001/83/EC, Regulation(EC)178/2002 and Regulation (EC) No 1223/2009 and repealing Council Directives 90/385/EEC and 93/42/EEC (Text with EEA relevance)*. Retrieved from https://op.europa.eu/en/publication-detail/-/publication/c262459f-bcb4-11ef-91ed-01aa75ed71a1

Valkonen, T. 1958-, & Lassila, J. (2021). *Väestön ikääntymisen taloudelliset vaikutukset*. Retrieved from https://julkaisut.valtioneuvosto.fi/handle/10024/163134

van Erven, T., & Harremoës, P. (2012). R\'enyi Divergence and Kullback-Leibler Divergence. *IEEE Transactions on Information Theory*, *60*(7), 3797–3820. doi: 10.1109/TIT.2014.2320500

Vehko, T., Hyppönen, H., Ryhänen, M., Tuukkanen, J., Ketola, E., & Heponiemi, T. (2018). Tietojärjestelmät ja työhyvinvointi – terveydenhuollon ammattilaisten näkemyksiä. *Finnish Journal of EHealth and EWelfare*, *10*(1), 143–163. doi: 10.23996/FJHW.65387

Wang, G., Li, W., Ourselin, S., & Vercauteren, T. (2019). Automatic Brain Tumor Segmentation Based on Cascaded Convolutional Neural Networks With Uncertainty Estimation. *Frontiers in Computational Neuroscience*, *13*, 56. doi: 10.3389/FNCOM.2019.00056

Wang, Y. X., Balle, B., & Kasiviswanathan, S. P. (2018). Subsampled R\'enyi Differential Privacy and Analytical Moments Accountant. *AISTATS 2019 - 22nd International Conference on Artificial Intelligence and Statistics*. doi: 10.29012/jpc.723

Weisstein, E. W. (n.d.). *L^2-Norm*. Retrieved from https://mathworld.wolfram.com/L2-Norm.html

Weng, W., & Zhu, X. (2015). U-Net: Convolutional Networks for Biomedical Image Segmentation. *IEEE Access*, *9*, 16591–16603. doi: 10.1109/ACCESS.2021.3053408

Zhou, T., Canu, S., Vera, P., & Ruan, S. (2021). 3D Medical Multi-modal Segmentation Network Guided by Multi-source Correlation Constraint. *Proceedings - International Conference on Pattern Recognition*, 10243–10250. doi: 10.1109/ICPR48806.2021.9412796