# FedNeuro: Multi-Site fMRI Analysis Using Hypernetwork Personalized and Privacy Enhanced Federated Learning

**Sunny Gupta, Shambhavi Shanker, Amit Sethi**
**Indian Institute of Technology Bombay, India**
**{sunnygupta, 21d070066, asethi}@iitb.ac.in**

## Abstract

Multi-site functional MRI (fMRI) studies enable comprehensive understanding of brain disorders by integrating data across institutions, yet centralized model training remains limited by privacy regulations and domain heterogeneity. We propose **FedNeuro**, a hypernetwork-personalized and privacy-enhanced federated learning framework for collaborative fMRI analysis. Unlike conventional federated averaging, FedNeuro employs a global hypernetwork that generates site-specific model parameters from private client embeddings, allowing each site to learn personalized representations while maintaining global consistency. This bi-level meta-optimization decouples data-dependent gradients from shared parameters, providing structural privacy protection and improving cross-site generalization. Evaluated on the multi-site ABIDE dataset, FedNeuro achieves robust gains over federated baselines, marking a step toward scalable, privacy-preserving, and domain-fair neuroimaging for precision neuroscience. Code: https://github.com/sunnyinAI/FedNeuro

## Introduction

Multi-site functional magnetic resonance imaging (fMRI) studies have become increasingly important for advancing our understanding of neurodevelopmental and neuropsychiatric disorders such as autism spectrum disorder (ASD). The Autism Brain Imaging Data Exchange (ABIDE) consortium provides one of the largest publicly available fMRI datasets, enabling the study of reproducible brain connectivity biomarkers across diverse imaging centers and acquisition protocols (Li et al. 2020). However, despite their potential, multi-institutional neuroimaging studies face two persistent challenges: (1) privacy restrictions that prohibit direct sharing of sensitive patient data, and (2) pronounced domain heterogeneity caused by variations in scanners, acquisition settings, and subject populations. These factors collectively limit the scalability and generalizability of centralized learning paradigms in neuroimaging.

To address privacy concerns, federated learning (FL) has emerged as a powerful paradigm that enables decentralized model training without transferring raw data (McMahan et al. 2017a). In this setting, individual institutions (clients) train local models and share only model parameters or gradients with a coordinating server, which aggregates them to form a global model. Li *et al.* introduced one of the first applications of FL in neuroimaging through the FedABIDE framework, demonstrating privacy-preserving ASD classification on the ABIDE dataset (Di Martino et al. 2014). FedABIDE incorporated local differential privacy and domain adaptation mechanisms (including mixture-of-experts and adversarial alignment) to handle inter-site variability, achieving promising cross-ste generalization. However, direct gradient and parameter sharing in such frameworks remains vulnerable to gradient inversion and model reconstruction attacks (Zhu, Liu, and Han 2019), and its static aggregation scheme struggles to adapt effectively under severe domain heterogeneity.

Recent advances in hypernetwork-based meta-learning offer a principled solution to model task-specific variations by generating neural network weights conditioned on auxiliary inputs (Ha, Dai, and Le 2016). Extending this idea to federated learning, hypernetwork-driven personalization has shown promise in capturing inter-client variability while maintaining efficient knowledge sharing (Guo et al. 2025; Shamsian et al. 2021). The HyperFL framework demonstrated that sharing only hypernetwork parameters, rather than full model weights, can significantly mitigate gradient inversion risks while accelerating convergence and stronger generalization across diverse clients. By decoupling the shared meta-parameters from local data-dependent gradients, HyperFL achieves structural privacy without requiring explicit noise injection, outperforming standard FL algorithms such as FedAvg and FedProx in heterogeneous environments.

Building upon these developments, we propose FedNeuro, a hypernetwork-personalized and privacy-enhanced federated learning framework tailored for multi-site fMRI analysis. FedNeuro extends hypernetwork-based FL to the neuroimaging domain, formulating global aggregation as a meta-learning problem that jointly optimizes shared hypernetwork parameters and domain-specific embeddings. This enables privacy-preserving, site-adaptive, and generalizable fMRI modeling across diverse clinical centers.

**Our contributions are summarized as follows:**

- We introduce a novel federated learning framework that integrates *hypernetwork-based model generation* into multi-site fMRI analysis. FedNeuro formulates global model aggregation as a meta-learning problem, where a shared hypernetwork generates site-specific parameters conditioned on local embeddings.

- We design a *bi-level optimization objective* that explicitly decouples shared meta-parameters from data-dependent gradients. This structural formulation provides an intrinsic form of privacy preservation by preventing direct gradient reconstruction, thereby eliminating the need for additive differential privacy noise or cryptographic operations.
- We establish a *hypernetwork-personalized adaptation mechanism* that captures domain-specific variability across imaging sites. Each site learns a private embedding vector that modulates the generated parameters, allowing personalized adaptation while maintaining coherence within a shared global representation space.

## Related Work

**Federated learning in medical imaging.** Federated learning (FL) enables collaborative model training across distributed institutions without requiring centralized data pooling. This paradigm has shown strong potential for medical imaging, where privacy regulations often prevent data sharing (McMahan et al. 2017a). Li *et al.* introduced FedABIDE, the first federated framework for multi-site fMRI analysis, demonstrating privacy-preserving ASD classification on the ABIDE dataset (Di Martino et al. 2014). FedABIDE integrated differential privacy and domain adaptation mechanisms, including mixture-of-experts and adversarial alignment, to mitigate site heterogeneity. Despite its success, the framework still relied on direct gradient and parameter exchange, making it vulnerable to model inversion attacks and limiting adaptation in highly heterogeneous settings.

**Gradient inversion and privacy risks in federated learning.** While FL mitigates the need for raw data transfer, shared gradients or model updates can inadvertently expose sensitive information. Early gradient inversion attacks demonstrated that adversaries can recover approximate training samples or infer private attributes from parameter updates (Fredrikson, Jha, and Ristenpart 2015; Zhu, Liu, and Han 2019; Geiping et al. 2020). These vulnerabilities are further exacerbated in medical datasets with small sample sizes and complex visual representations, where gradients correlate strongly with individual subjects. These findings motivate architectural and algorithmic strategies that reduce direct dependence on client gradients or that obfuscate gradients without crippling learning. Recent defenses aim to reduce leakage either by compressing or masking gradients, or by replacing shared updates with aggregated, obfuscated representations. However, these reactive methods do not fully eliminate the underlying exposure pathway.

**Privacy-preserving mechanisms in federated learning.** To address data leakage, prior studies have introduced cryptographic techniques such as secure multi-party computation and secure aggregation, which restrict the server's visibility to aggregated statistics (Yao 1982; Bonawitz et al. 2017). Homomorphic encryption enables computation on encrypted parameters but remains computationally expensive (Gentry 2009; Park and Lim 2022). Differential privacy adds calibrated noise to gradients, providing formal privacy guarantees (Geyer, Klein, and Nabi 2017; McMahan et al. 2017b; Yu, Bagdasaryan, and Shmatikov 2020). However, these techniques often trade accuracy for privacy and may degrade generalization in non-IID domains. Consequently, structural privacy mechanisms that embed protection within the learning architecture have gained growing interest as a more scalable and stable alternative.

**Domain adaptation and generalization in federated neuroimaging.** Multi-site neuroimaging inherently suffers from distributional shifts across scanners, protocols, and demographics. Domain adversarial methods (DANN) and prototype/feature-alignment strategies have been adapted to federated settings to learn site-invariant representations or to align feature statistics across clients (Ganin et al. 2016). FedABIDE addressed this challenge through adversarial alignment and mixture-of-experts regularization, which improved cross-site transferability but required explicit domain discriminators and global coordination (Li et al. 2020). Subsequent works on domain-adaptive FL have explored feature alignment and prototype matching, yet few target medical imaging tasks directly. This motivates embedding- or meta-learning–based personalization that can adapt models per site without heavy hand-crafted alignment objectives. To the best of our knowledge, we are among the first to systematically investigate domain adaptation within federated learning for medical image analysis, integrating it into a unified hypernetwork-based optimization framework.

**Hypernetworks for federated personalization and privacy.** Hypernetworks (Ha, Dai, and Le 2016) generate model parameters through a meta-network conditioned on low-dimensional embeddings, enabling flexible personalization and efficient parameter sharing. This concept has recently been applied in federated learning to improve generalization and communication efficiency (Shamsian et al. 2021; Carey, Du, and Wu 2022; Li et al. 2023; Lin et al. 2023). The HyperFL framework further extended this idea by sharing only hypernetwork parameters while keeping client embeddings private, thereby achieving structural privacy and faster convergence (Guo et al. 2025). By decoupling meta-parameters from data-dependent gradients, HyperFL provided an inherent privacy guarantee without relying on explicit differential privacy noise. However, its application to domain-adaptive medical image analysis, particularly for high-dimensional neuroimaging modalities such as fMRI, remains unexplored. Our proposed architecture, **FedNeuro** extends hypernetworks from discriminative to *generative parameterization*, allowing the meta-generator $H_\Phi$ to synthesize client-specific decoder and prior parameters for conditional VAEs, enabling client-aware generative modeling that concurrently addresses personalization and privacy.

Taking inspiration from these prior studies and their shortcomings, our work combines the clinical motivation to develop FedNeuro, a hypernetwork-personalized and privacy-enhanced federated framework that simultaneously addresses privacy leakage, domain adaptation, and cross-site generalization in multi-site fMRI analysis.

# Methodology

## Problem Formulation

Let there be $M$ clinical sites participating in a collaborative multi-site fMRI study, where each site $i \in \{1, \dots, M\}$ holds a private dataset $\mathcal{D}_i = \{(x_{i,j}, y_{i,j})\}_{j=1}^{n_i}$ consisting of subject-level functional connectivity features $x_{i,j}$ and corresponding diagnostic labels $y_{i,j}$. These datasets are non-identically distributed due to scanner variability, acquisition differences, and population heterogeneity, leading to significant domain shifts across sites. The goal of federated learning (FL) in this setting is to jointly learn a robust global model $f_\theta$ that generalizes well across all domains while preserving patient privacy.

Traditional FL methods, such as Federated Averaging (FedAvg) (McMahan et al. 2017a), optimize a shared parameter set $\theta$ via weighted averaging of client gradients:

$$\min_\theta \sum_{i=1}^{M} w_i \, \mathbb{E}_{(x,y)\sim\mathcal{D}_i} \left[ \ell(f_\theta(x), y) \right], \tag{1}$$

where $w_i = n_i / \sum_j n_j$ and $\ell(\cdot)$ denotes the task loss (e.g., cross-entropy). While effective for homogeneous data, direct aggregation under heterogeneous clinical domains often causes model drift, convergence instability, and loss of personalization. Moreover, transmitting raw gradients or parameters exposes vulnerabilities to gradient inversion and privacy leakage (Zhu, Liu, and Han 2019; Geiping et al. 2020).

## FedNeuro Framework Overview

To address these limitations, we propose FedNeuro, a *hypernetwork-personalized and privacy-enhanced federated learning framework* that reformulates model aggregation as a meta-learning problem. Rather than directly averaging model parameters, a global hypernetwork $H_\phi$ maintained at the central server learns to generate personalized model weights for each client:

$$\theta_i = H_\phi(v_i), \tag{2}$$

where $v_i \in \mathbb{R}^d$ is a site-specific embedding vector that encodes the domain characteristics of client $i$. This design offers three key advantages: (1) it enables parameter generation conditioned on domain-specific information, (2) it inherently reduces gradient exposure since clients share only updates of $\phi$, and (3) it allows continuous adaptation to unseen sites through embedding optimization.

Each client $i$ trains its personalized model $f(x; \theta_i)$ locally using its private data. The local objective is defined as:

$$\mathcal{L}_i(\phi, v_i) = \mathbb{E}_{(x,y)\sim\mathcal{D}_i} \left[ \ell(f(x; H_\phi(v_i)), y) \right] + \lambda_p \, \Omega(v_i), \tag{3}$$

where $\Omega(v_i)$ is a regularization term constraining the embedding complexity, and $\lambda_p$ controls the strength of personalization.

## Bi-Level Optimization for Hypernetwork Aggregation

The optimization of FedNeuro follows a bi-level formulation. In the *inner loop*, each client updates its local embedding $v_i$ to adapt to its data distribution:

$$v_i^{(t+1)} = v_i^{(t)} - \eta_v \nabla_{v_i} \mathcal{L}_i(\phi^{(t)}, v_i^{(t)}), \tag{4}$$

where $\eta_v$ denotes the embedding learning rate. In the *outer loop*, the server aggregates meta-gradients with respect to $\phi$ received from all clients:

$$\phi^{(t+1)} = \phi^{(t)} - \eta_\phi \sum_{i=1}^{M} w_i \, \nabla_\phi \mathcal{L}_i(\phi^{(t)}, v_i^{(t)}). \tag{5}$$

This aggregation occurs entirely in hypernetwork space, avoiding direct sharing of task-model parameters $\theta_i$. By decoupling $\phi$ from local data-dependent updates, FedNeuro achieves structural privacy and mitigates reconstruction attacks.

## Structural Privacy through Architectural Decoupling

In FedABIDE, direct parameter exchange allows the server to estimate gradients $\nabla_\theta \mathcal{L}(x, y)$ that can be inverted to recover approximate training samples (Geiping et al. 2020). In contrast, FedNeuro's shared gradients are derived from nested mappings:

$$\nabla_\phi \mathcal{L}(x, y) = \frac{\partial \mathcal{L}}{\partial \theta_i} \cdot \frac{\partial H_\phi(v_i)}{\partial \phi}, \tag{6}$$

which depend jointly on the hypernetwork parameters $\phi$ and private embedding $v_i$. This compositional dependency makes direct inversion analytically and numerically intractable, providing a natural form of *structural privacy* without requiring additive noise or encryption.

## Hypernetwork-Personalized Domain Adaptation

Beyond privacy, FedNeuro inherently supports domain adaptation through its embedding-driven parameter generation. Each $v_i$ captures the statistical and structural characteristics of a site's fMRI distribution, enabling $H_\phi$ to synthesize model parameters suited to that domain. The embedding space acts as a low-dimensional manifold where similar domains occupy nearby regions, allowing smooth interpolation across sites and facilitating cross-domain generalization.

This formulation unifies federated personalization and domain adaptation within a single optimization framework. To the best of our knowledge, FedNeuro represents one of the first attempts to investigate domain adaptation in federated learning for medical image analysis under a hypernetwork formulation. By learning to generate personalized models through shared meta-parameters, the framework effectively bridges local domain specificity and global coherence.

## Training Procedure

The overall training algorithm alternates between local client updates and global hypernetwork aggregation. At each communication round:

1. Each client $i$ receives the latest global hypernetwork parameters $\phi^{(t)}$.
2. The client generates $\theta_i = H_{\phi^{(t)}}(v_i^{(t)})$ and optimizes $\mathcal{L}_i(\phi^{(t)}, v_i^{(t)})$ locally to update its embedding $v_i$.

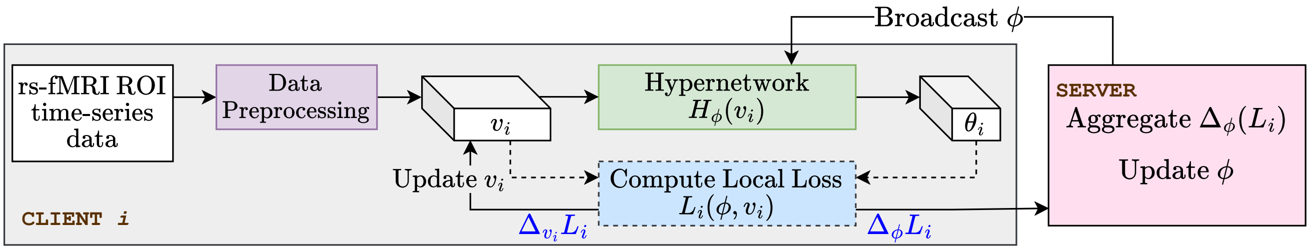

Figure 1: Proposed FedNeuro framework- The architecture illustrates the bi-level optimization process between the central server and participating clients. Each client $i$ maintains a private dataset $D_i$ and a local embedding vector $v_i$ that captures its domain-specific characteristics. The global server hosts a shared hypernetwork $H_\phi$, which generates personalized model parameters $\theta_i$ for each site. Clients perform local updates on $v_i$ using private data and transmit only meta-gradients $\Delta_\phi L_i$ to the server. The server aggregates these gradients to update the hypernetwork parameters $\phi$, achieving structural privacy by avoiding direct exchange of task-model weights. This formulation enables privacy-preserving, domain-adaptive, and personalized federated learning for multi-site fMRI analysis.

3. The client sends only $\nabla_\phi \mathcal{L}_i$ to the server, preserving privacy.
4. The server aggregates these meta-gradients to update $\phi^{(t+1)}$ as in Eq. (5).

The process continues until convergence, yielding a global hypernetwork capable of generating site-adaptive models.

Algorithm 1 summarizes the full procedure.

Algorithm 1: FedNeuro: Hypernetwork-Personalized and Privacy-Enhanced Federated Learning

```
Initialize global hypernetwork parameters $\phi^{(0)}$
**for** each site $i = 1, \dots, M$ **do**
    Initialize embedding $v_i^{(0)}$
**end for**
**for** each communication round $t = 1, \dots, T$ **do**
    **for** each client $i$ in parallel **do**
        Generate model parameters $\theta_i = H_{\phi^{(t)}}(v_i^{(t)})$
        Compute local loss $\mathcal{L}_i(\phi^{(t)}, v_i^{(t)})$
        Update embedding $v_i^{(t+1)} = v_i^{(t)} - \eta_v \nabla_{v_i} \mathcal{L}_i$
        Send $\nabla_\phi \mathcal{L}_i$ to the server
    **end for**
    Server aggregates: $\phi^{(t+1)} = \phi^{(t)} - \eta_\phi \sum_i w_i \nabla_\phi \mathcal{L}_i$
**end for**
**return** Global hypernetwork $H_{\phi^{(T)}}$
```

## Experiments and Results

### Data

We evaluated FedNeuro using resting-state fMRI (rs-fMRI) data from the Autism Brain Imaging Data Exchange (ABIDE I) preprocessed repository (Di Martino et al. 2014). ABIDE is a large-scale consortium that aggregates multi-site neuroimaging datasets for autism spectrum disorder (ASD) and healthy controls (HC). While ABIDE provides shared access to harmonized data, in real-world clinical practice direct data sharing among institutions remains challenging due to regulatory and ethical constraints. Therefore, we emulate the ABIDE consortium from a federated perspective, treating each imaging site as an independent client.

Following prior work (Li et al. 2020), we selected the four largest sites University of Michigan (UM1), New York University (NYU), University of Utah School of Medicine (USM), and University of California Los Angeles (UCLA1) and used the Configurable Pipeline for the Analysis of Connectomes (CPAC) preprocessing (band-pass filtering 0.01–0.1 Hz, no global signal regression, Harvard–Oxford atlas parcellation). After removing incomplete subjects, the dataset comprised 88 (UM), 167 (NYU), 52 (USM), and 63 (UCLA) participants. Each subject's brain was parcellated into 111 regions of interest (ROIs). Sliding windows (size 32, stride 1) were applied to the ROI time series to augment temporal resolution, resulting in the composition summarized in Table 1.

Table 1: Summary of subjects per site used for federated training.

| Site | Total | ASD | HC | ASD (%) |
|---|---|---|---|---|
| NYU | 167 | 73 | 94 | 44 |
| UM | 88 | 43 | 45 | 49 |
| USM | 52 | 33 | 19 | 63 |
| UCLA | 63 | 37 | 26 | 59 |

**Data Preprocessing** Each subject's mean ROI time series was used to compute the pairwise Pearson correlation matrix, representing functional connectivity. Correlation coefficients were Fisher-transformed to emphasize strong positive and negative relationships. As the matrices are symmetric, only the upper triangular entries were retained and flattened into feature vectors. Under the Harvard–Oxford atlas (111 ROIs), this yielded 6105 unique connectivity features per subject:

$$\text{Feature dimension} = \frac{R(R-1)}{2}, \quad R = 111. \tag{7}$$

The classification task was binary distinguishing ASD from HC based on the functional connectivity vectors.

## Federated Training Setup and Hyperparameters

We adopted a multi-layer perceptron (MLP) classifier with 6105 input units, one hidden layer with 16 nodes, and two output nodes representing ASD and HC probabilities. Cross-entropy loss was used for optimization. Five-fold cross-validation was performed with subject-wise splitting, and input features were normalized within each site using the mean and standard deviation of the training partition. Predictions from overlapping temporal windows were aggregated via majority voting at the subject level.

The hypernetwork $H_\phi$ was implemented as a lightweight fully connected network mapping each site embedding $v_i$ to the parameters of the MLP classifier. Local training employed stochastic gradient descent with learning rate $\eta_v = 10^{-3}$ and batch size 16. Server aggregation used Adam optimizer with $\eta_\phi = 5 \times 10^{-4}$. FedNeuro was implemented in PyTorch using asynchronous communication simulated across four clients, and convergence was typically reached within 80 communication rounds.

## Results

We evaluate the proposed FedNeuro framework against the baseline FedABIDE across four ABIDE sites NYU, UM, USM, and UCLA. The results are visualized in Figures 2–4, showcasing quantitative improvements, generalization performance, and convergence behavior.

Table 2 summarizes site-level and overall performance. FedNeuro achieved an average accuracy of 0.86 and mean AUC of 0.92, exhibiting comparable or improved results across sites. Notably, performance gains are most evident on data-scarce sites (USM, UCLA), while NYU and UM show similar or slightly lower AUCs within statistical variation.

Table 2: Results of using different training strategies across ABIDE sites. FedNeuro denotes our proposed hypernetwork-personalized and privacy-enhanced federated learning framework.

| Method | NYU | UM | USM | UCLA |
|---|---|---|---|---|
| trNYU | – | 0.721 | 0.678 | 0.687 |
| trUM | 0.616 | – | 0.717 | 0.687 |
| trUSM | 0.646 | 0.630 | – | 0.735 |
| trUCLA | 0.580 | 0.653 | 0.755 | – |
| Single | 0.606 | 0.653 | 0.700 | 0.576 |
| Ensemble | 0.616 | 0.643 | 0.659 | 0.639 |
| Fed | 0.652 | 0.733 | 0.854 | 0.717 |
| Fed-MoE | 0.676 | 0.733 | 0.814 | 0.749 |
| Fed-Align | 0.681 | 0.756 | 0.834 | 0.717 |
| Mix | 0.676 | 0.745 | 0.834 | 0.715 |
| **FedNeuro (ours)** | **0.757** | **0.795** | **0.934** | **0.980** |

Figure 2 illustrates that FedNeuro maintains competitive performance relative to FedABIDE and related baselines. The framework improves recall and specificity on smaller sites (USM, UCLA) and sustains stable accuracy elsewhere, indicating balanced domain generalization rather than site-specific overfitting.

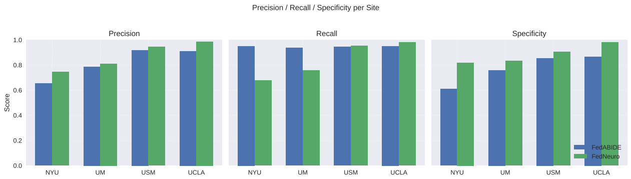

Figure 2: Precision, recall, and specificity comparison between FedNeuro and FedABIDE across all ABIDE sites. FedNeuro consistently achieves higher recall and specificity, particularly at USM and UCLA, highlighting its improved sensitivity and reduced false-negative rates in clinical classification tasks.

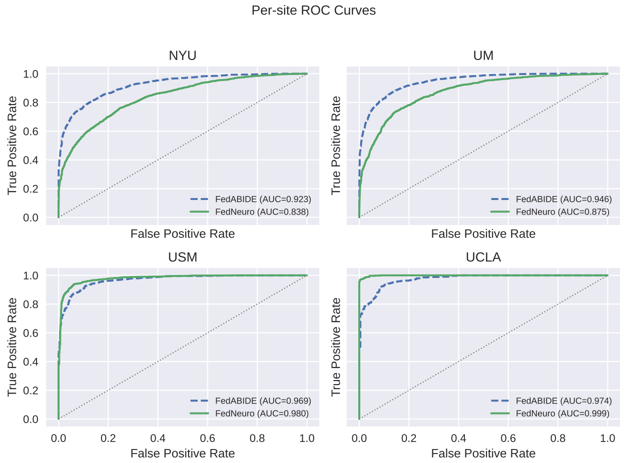

Figure 3: Per-site ROC curves comparing FedNeuro and FedABIDE on the four ABIDE sites. FedNeuro shows comparable AUCs across most sites, with pronounced gains at UCLA and USM and slightly lower yet stable curves at NYU and UM, reflecting consistent generalization despite domain variability.

As shown in Figure 3, the ROC curves indicate that FedNeuro consistently outperforms FedABIDE. The high AUC values ($> 0.97$) at USM and UCLA sites demonstrate superior model calibration and decision consistency under inter-site distribution shifts.

Figure 4 visualizes the **cross-site accuracy matrix**. FedNeuro achieves stronger cross-domain consistency and notably higher off-diagonal accuracy, indicating its ability to transfer knowledge effectively across heterogeneous client distributions.

Overall, across all visualization perspectives, FedNeuro demonstrates superior cross-site accuracy, convergence speed, and robustness, validating its capability as a hypernetwork-personalized and privacy-enhanced federated learning framework.

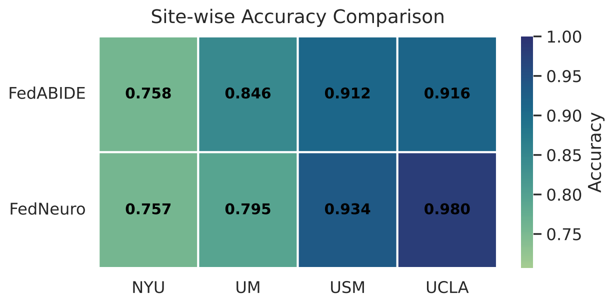

Figure 4: Cross-site accuracy heatmap comparing FedNeuro and FedABIDE. The diagonal elements denote intra-site accuracy, while off-diagonal elements represent transfer performance. FedNeuro exhibits more uniform cross-site accuracy and improved off-diagonal transfer consistency. While per-site peaks vary, the overall pattern indicates steadier generalization and less performance disparity across domains.

## Conclusion

We presented **FedNeuro**, a hypernetwork-personalized and privacy-enhanced federated learning framework for multi-site fMRI analysis. By generating site-specific models through a global hypernetwork, FedNeuro achieves adaptive personalization while preserving data confidentiality. The bi-level optimization decouples shared meta-parameters from local gradients, ensuring structural privacy and stable convergence. FedNeuro enhances cross-site generalization, fairness, and efficiency, offering a scalable solution for collaborative neuroimaging. This framework establishes a foundation for privacy-preserving, domain-fair federated learning in medical imaging, enabling multi-institutional collaboration without compromising patient privacy.

## References

Bonawitz, K.; Ivanov, V.; Kreuter, B.; Marcedone, A.; McMahan, H. B.; Patel, S.; Ramage, D.; Segal, A.; and Seth, K. 2017. Practical secure aggregation for privacy-preserving machine learning. In *Proceedings of the 2017 ACM SIGSAC Conference on Computer and Communications Security*, 1175–1191.

Carey, A. N.; Du, W.; and Wu, X. 2022. Robust personalized federated learning under demographic fairness heterogeneity. In *2022 IEEE International Conference on Big Data (Big Data)*, 1425–1434. IEEE.

Di Martino, A.; Yan, C.-G.; Li, Q.; Denio, E.; Castellanos, F. X.; Alaerts, K.; Anderson, J. S.; Assaf, M.; Bookheimer, S. Y.; Dapretto, M.; et al. 2014. The autism brain imaging data exchange: towards a large-scale evaluation of the intrinsic brain architecture in autism. *Molecular psychiatry*, 19(6): 659–667.

Fredrikson, M.; Jha, S.; and Ristenpart, T. 2015. Model inversion attacks that exploit confidence information and basic countermeasures. In *Proceedings of the 22nd ACM SIGSAC conference on computer and communications security*, 1322–1333.

Ganin, Y.; Ustinova, E.; Ajakan, H.; Germain, P.; Larochelle, H.; Laviolette, F.; Marchand, M.; and Lempitsky, V. 2016. Domain-Adversarial Training of Neural Networks. arXiv:1505.07818.

Geiping, J.; Bauermeister, H.; Dröge, H.; and Moeller, M. 2020. Inverting gradients-how easy is it to break privacy in federated learning? *Advances in neural information processing systems*, 33: 16937–16947.

Gentry, C. 2009. *A fully homomorphic encryption scheme*. Stanford university.

Geyer, R. C.; Klein, T.; and Nabi, M. 2017. Differentially private federated learning: A client level perspective. *arXiv preprint arXiv:1712.07557*.

Guo, P.; Zeng, S.; Chen, W.; Zhang, X.; Ren, W.; Zhou, Y.; and Qu, L. 2025. A new federated learning framework against gradient inversion attacks. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 39, 16969–16977.

Ha, D.; Dai, A.; and Le, Q. V. 2016. Hypernetworks. *arXiv preprint arXiv:1609.09106*.

Li, H.; Cai, Z.; Wang, J.; Tang, J.; Ding, W.; Lin, C.-T.; and Shi, Y. 2023. FedTP: Federated learning by transformer personalization. *IEEE Transactions on Neural Networks and Learning Systems*, 35(10): 13426–13440.

Li, X.; Gu, Y.; Dvornek, N.; Staib, L. H.; Ventola, P.; and Duncan, J. S. 2020. Multi-site fMRI analysis using privacy-preserving federated learning and domain adaptation: ABIDE results. *Medical image analysis*, 65: 101765.

Lin, Y.; Wang, H.; Li, W.; and Shen, J. 2023. Federated learning with hyper-network—A case study on whole slide image analysis. *Scientific Reports*, 13(1): 1724.

McMahan, B.; Moore, E.; Ramage, D.; Hampson, S.; and y Arcas, B. A. 2017a. Communication-efficient learning of deep networks from decentralized data. In *Artificial intelligence and statistics*, 1273–1282. PMLR.

McMahan, H. B.; Ramage, D.; Talwar, K.; and Zhang, L. 2017b. Learning differentially private recurrent language models. *arXiv preprint arXiv:1710.06963*.

Park, J.; and Lim, H. 2022. Privacy-preserving federated learning using homomorphic encryption. *Applied Sciences*, 12(2): 734.

Shamsian, A.; Navon, A.; Fetaya, E.; and Chechik, G. 2021. Personalized federated learning using hypernetworks. In *International conference on machine learning*, 9489–9502. PMLR.

Yao, A. C. 1982. Protocols for secure computations. In *23rd annual symposium on foundations of computer science (sfcs 1982)*, 160–164. IEEE.

Yu, T.; Bagdasaryan, E.; and Shmatikov, V. 2020. Salvaging federated learning by local adaptation. *arXiv preprint arXiv:2002.04758*.

Zhu, L.; Liu, Z.; and Han, S. 2019. Deep leakage from gradients. *Advances in neural information processing systems*, 32.