BreastG-FCL: Graph-Conditioned Federated Continual Learning for Breast Cancer Radiogenomics

Qingyang Yu
Electrical and Computer Engineering
Stevens Institute of Technology
Hoboken, United States
qyu13@stevens.edu

Hao Wang
Electrical and Computer Engineering
Stevens Institute of Technology
Hoboken, United States
hwang9@stevens.edu

Jingyi Wang
Computer Science
San Francisco State University
San Francisco, United States
jingyiwang@sfsu.edu

Xinyue Zhang
Computer Science
Kennesaw State University
Kennesaw, United States
xzhang48@kennesaw.edu

Miao Pan
Electrical and Computer Engineering
University of Houston
Houston, United States
mpan2@uh.edu

Ziyue Xu
NVIDIA
Bethesda, United States
ziyuex@nvidia.com

***Abstract*—Breast cancer radiogenomics has been widely studied using machine learning to uncover associations between imaging phenotypes and molecular characteristics. However, real-world deployment requires privacy-preserving learning across heterogeneous institutions and continual adaptation to evolving clinical distributions, challenges that conventional models trained on static and limited datasets cannot adequately address. Federated continual learning (FCL) offers an effective solution by enabling collaborative learning and continual adaptation while keeping raw data with privacy preservation. Nevertheless, existing FCL methods mainly rely on generic signals, such as whole-image representations or model updates, and thus overlook fine-grained disease-specific cues, including tumor morphology and dynamic contrast-enhanced MRI (DCE-MRI) kinetics. To address this limitation, we propose *BreastG-FCL*, a disease-aware, graph-conditioned FCL framework for breast cancer radiogenomic modeling. *BreastG-FCL* constructs relational graphs from tumor morphology and DCE-MRI kinetic patterns and uses them to guide generative replay and representation alignment across clients and tasks. Experiments on *TCGA-BRCA* show that *BreastG-FCL* achieves an accuracy of 84.87%, outperforming the best-performing baseline by 9.55 percentage points.**

***Index Terms*—Federated continual learning, radiogenomic modeling, breast cancer.**

I. Introduction

Breast cancer radiogenomics links imaging phenotypes with molecular characteristics to enable noninvasive prediction of clinical outcomes, including tumor subtypes [1]. Researchers have developed numerous machine learning (ML) methods to improve the accuracy of these associations [2]. However, patient heterogeneity requires large and diverse training datasets that individual institutions often cannot assemble [3]. In addition, breast cancer radiogenomics involves evolving data distributions, including dynamic contrast-enhanced MRI (DCE-MRI) sequences that capture temporal contrast-enhancement patterns [4]. Conventional ML models, such as ResNet [5], assume static training data and cannot continually adapt to evolving data distributions without forgetting previously learned knowledge. Consequently, individual institutions struggle to train models that generalize across diverse patients while adapting to newly arriving data. To address both challenges while complying with privacy regulations such as HIPAA [6], prior studies integrate federated learning (FL), which enables privacy-preserving collaborative training [7], with continual learning (CL), which supports continual adaptation to sequential data [8], thereby forming federated continual learning (FCL).

Accordingly, a growing body of FCL studies has emerged. Regularization-based methods constrain model drift across tasks [9], while task-based methods leverage task information to facilitate adaptation [3]. Replay-based methods generate synthetic data or representations for previous tasks and have become a common choice for FCL [10]. Other recent studies further explore through a mathematical aspect, such as analytic solutions [11]. Among these methods, the prior study [12] constructs spatiotemporal relational graphs from client model updates and uses them to guide representation alignment, achieving state-of-the-art (SOTA) performance. However, model update-derived graphs encode spatial relations based on generic data-distribution similarities rather than disease-specific radiogenomic regions, while temporal relations merely reflect changes in these spatial relations.

Such a limitation is critical in breast cancer radiogenomics, where effective modeling requires fine-grained, disease-specific relations. As shown in Figure 1, in breast MRI, clinically meaningful spatial relations arise from localized regions, including the tumor core, enhancing rim, peritumoral tissue, and background parenchymal enhancement (BPE) [13], rather than generic data distributions. Figure 2 shows that, in DCE-MRI, temporal relations are characterized by contrast-enhancement kinetics rather than naive changes in spatial attention.

We therefore propose *BreastG-FCL* for breast cancer radiogenomic modeling [14], a graph-conditioned FCL framework. Its spatial relations model disease-relevant breast regions, while temporal relations use wash-in, wash-out, time-to-peak (TTP),

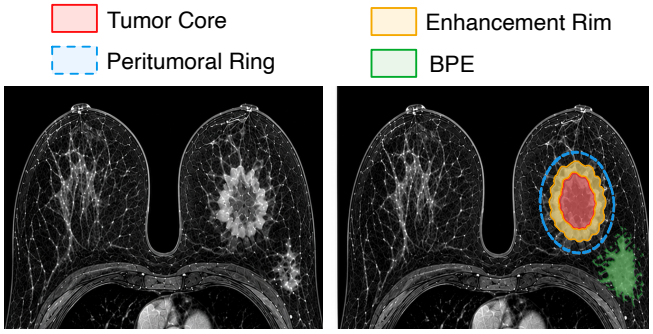

Fig. 1: Spatial relations in breast cancer radiogenomics should be constructed from disease-specific regions rather than the whole magnetic resonance imaging (MRI) image. **Left:** Original breast MRI without regional annotations. **Right:** The MRI with highlighted disease-specific regions.

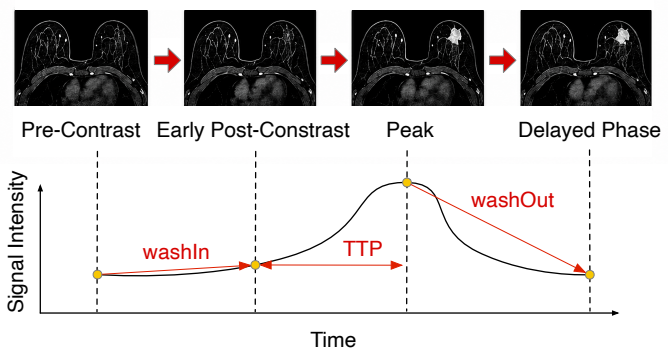

Fig. 2: DCE-MRI enhancement pattern used for kinetics-aware temporal relations. The upper row illustrates the sequential imaging phases from pre-contrast to early post-contrast, peak enhancement, and delayed washout. The lower plot shows the corresponding signal-intensity curve.

and signal enhancement ratio (SER) [4] to model temporal patterns in DCE-MRI. These clinical cues capture inter-client relations that generic model update-derived graphs cannot represent.

Our contributions are summarized as follows:

- We propose *BreastG-FCL*, a fine-grained spatial-temporal graph-conditioned FCL framework tailored to breast cancer radiogenomic modeling.
- We construct disease-aware spatial relations from regional tumor morphology and temporal relations from DCE-MRI enhancement kinetics.
- We conduct extensive experiments on *TCGA-BRCA* [15] to evaluate *BreastG-FCL* against representative FCL baselines. Results show that *BreastG-FCL* improves breast cancer radiogenomic prediction and strengthens adaptation to healthcare non-IID drift under federated privacy constraints with an improvement of 9.55 percentage points in accuracy.

II. Background and Motivation

A. Breast Cancer Radiogenomics

Breast cancer radiogenomics links imaging phenotypes with molecular characteristics to support the noninvasive prediction of tumor subtypes and clinical outcomes [1]. The matched *TCGA-BRCA* and *TCIA* [16] data reveal complementary spatial and temporal heterogeneity. Spatially, disease-relevant information is concentrated in the tumor core, enhancing rim, peritumoral tissue, and BPE, whose morphology reflects tumor and microenvironmental characteristics [13]. Temporally, DCE-MRI captures contrast-enhancement trajectories characterized by wash-in, wash-out, TTP, and SER [4]. Together with gene-expression and clinical data, these observations motivate joint modeling of regional morphology and enhancement kinetics to capture clinically meaningful inter-patient heterogeneity.

B. Federated Continual Learning

In FCL, each of N clients C_i, $i \in [N]$, learns a sequence of K tasks $\{T_i^k\}_{k\in[K]}$. Task T_i^k provides a labeled dataset

$$\mathcal{D}_i^k = \{(x_{i,j}^k, y_{i,j}^k)\}_{j=1}^{n_i^k}, \quad (x_{i,j}^k, y_{i,j}^k) \sim T_i^k,$$

where n_i^k denotes the sample size of task T_i^k. When client C_i trains on the current task T_i^k, privacy and storage constraints prevent it from retaining or revisiting raw data from previous tasks $\{T_i^t\}_{t<k}$, *i.e.*, $\{\mathcal{D}_i^t\}_{t<k}$. Thus, client C_i must preserve prior-task knowledge through its model state and permitted replay or summary signals rather than raw historical samples. The training process aims to maximize downstream performance, *e.g.*, classification accuracy, across all clients and tasks.

Existing FCL methods broadly include regularization-based approaches [9], task-based transfer [3], replay-based methods [10], and others [11]. Among all these categories, replay-based methods are more competitive due to the use of generative models, *e.g.*, generative adversarial network (GAN), to replay synthetic historical data to improve training [8]. However, most replay-based FCL methods condition generation on labels, overlooking auxiliary information introduced by FCL scenarios beneficial to replay, *e.g.*, spatial and temporal relations among clients across tasks, thereby achieving suboptimal performance [17]. Thereby, the prior study [12] utilizes relational graphs to capture the spatial and temporal relations among clients' data distributions in FCL scenarios, and conditions generation on relational graphs, achieving SOTA performance.

C. Generic vs. Disease-Specific Spatial and Temporal Relations

Although Yu et al. [12] uses relational graphs to capture spatial and temporal relations among clients' data distributions in FCL scenarios, this generic graph construction misses disease-specific structure in breast cancer radiogenomic modeling. As shown in Figure 1, the inputs for spatial relations in generic graph construction are whole-MRI images, whereas in breast cancer radiogenomics, *e.g.*, *TCGA-BRCA*, clinically meaningful spatial relations instead originate from localized, disease-relevant regions, including tumor core, enhancing rim, peritumoral tissue, and BPE.

Similarly, generic temporal attention constructs temporal relations by concatenating or tracking spatial attention patterns over time [12]. However, as shown in Figure 2, in the *TCIA* imaging subset matched to *TCGA-BRCA*, DCE-MRI provides disease-specific temporal cues through contrast-enhancement kinetics, capturing breast cancer enhancement dynamics through wash-in, wash-out, TTP, and SER. Similar conclusions have been drawn by previous studies, which show that disease-specific regions contain complementary biological information obscured by the whole-MRI images [4]. This motivation can be further explained by the information bottleneck in information theory [18]. Specifically, denote $x_{\text{full}} = (x_{\text{fine}}, x_{\text{irrelevant}})$ as the full-size MRI and x_{fine} as the disease-specific regions in MRI, then the disease-specific graph construction can be interpreted as a task-oriented information bottleneck constrained by clinical priors, which keeps information relevant to the tumor, *e.g.*, x_{fine}, but compresses most irrelevant information to the tumor, $x_{\text{irrelevant}}$, *e.g.*, field strength and registration errors:

$$\max_{x_{\text{fine}}} I(x_{\text{fine}}; y) - \gamma I(x_{\text{fine}}; x_{\text{full}}).$$

Therefore, we customize spatial and temporal relations in the prior study [12] for breast cancer radiogenomics, *TCGA-BRCA*. Specifically, we construct spatial relations from disease-specific MRI regions and temporal relations from DCE-MRI enhancement curves. These disease-aware relational signals allow the framework to better capture clinical relations and strengthen radiogenomic prediction under FCL scenarios.

III. System Design

A. Overview

BreastG-FCL follows the FCL architecture in the prior study [12], where one global discriminator D is initialized on the central server and local encoder-generator-predictor pairs initialized in N clients. Each client C_i keeps raw data locally and learns a sequence of Tasks $\{T_i^k\}_{k=1}^K$. At Task k, client C_i trains on the current dataset $\mathcal{D}_i^k$ without accessing raw historical datasets $\{\mathcal{D}_i^j\}_{j<k}$.

Formally, for Task k, *BreastG-FCL* performs a minimax optimization with the value function $V(E_i, F_i, \mathcal{G}_i, D)$ as:

$$\min_{\{E_i,\mathcal{G}_i,F_i\}_{i=1}^N} \max_D \sum_{i=1}^N [V_{p,i}^k(E_i, F_i) - \lambda_d V_{d,i}^k(D, E_i) + \sum_{j=1}^k \left(V_{p,i}^{j,\text{syn}}(\mathcal{G}_i, F_i) - \lambda_d V_{d,i}^{j,\text{syn}}(D, \mathcal{G}_i)\right)],$$

where

$$V_{p,i}^k(E_i, F_i) = \mathbb{E}_{(x,y)\sim T_i^k}\left[\ell_p\big(F_i(E_i(x, g_i^k)), y\big)\right],$$
$$V_{d,i}^k(D, E_i) = \mathbb{E}_{(x,y)\sim T_i^k}\left[\ell_d\big(D(E_i(x, g_i^k) + \tau_i), \hat{g}_i^k\big)\right],$$
$$V_{p,i}^{j,\text{syn}}(\mathcal{G}_i, F_i) = \mathbb{E}_{z,\, y\sim T_i^j}\left[\ell_p\big(F_i(\mathcal{G}_i(z, y, g_i^j)), y\big)\right],$$
$$V_{d,i}^{j,\text{syn}}(D, \mathcal{G}_i) = \mathbb{E}_{z,\, y\sim T_i^j}\left[\ell_d\big(D(\mathcal{G}_i(z, y, g_i^j) + \tau_i), \hat{g}_i^j\big)\right].$$

Here, $z \sim \mathcal{N}(0, I)$ is a random Gaussian noise, $\tau_g, \tau_i \sim \text{Lap}(0, 1)$ are Laplace noise terms introduced for privacy protection, $\hat{g}_i^j = g_i^j + \tau_g$ denotes the graph row added with the noise, ℓ_p denotes the prediction loss, and ℓ_d denotes the discriminator loss.

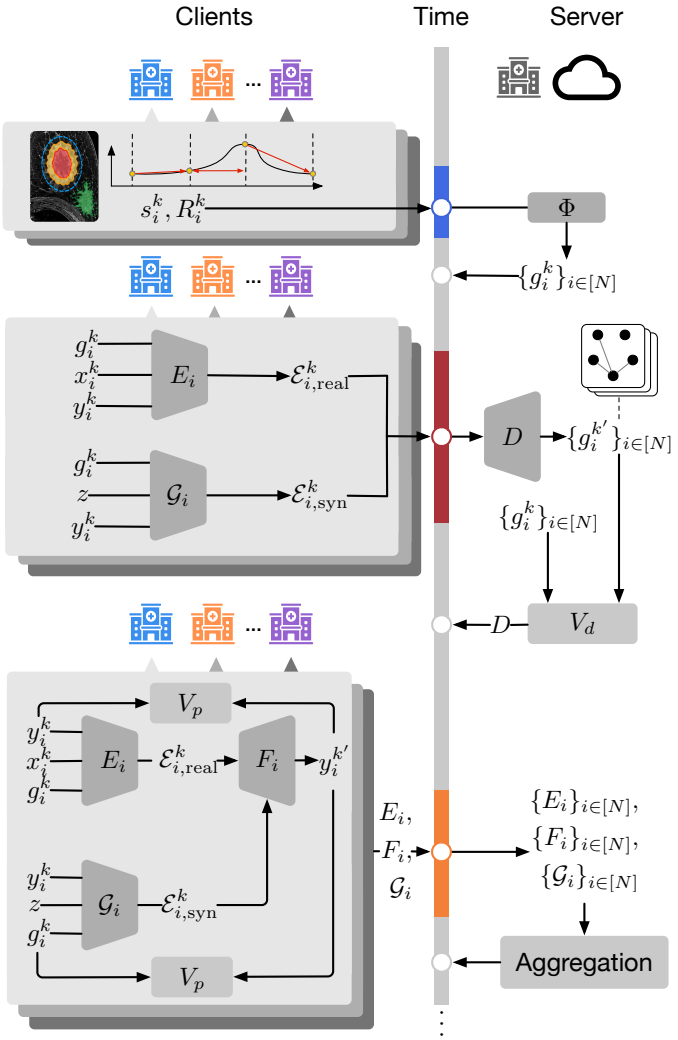

Fig. 3: Workflow of *BreastG-FCL* for Task k.

The workflow of *BreastG-FCL* for Task k is presented in Figure 3. Specifically, within Task k, *BreastG-FCL* decomposes the workflow into the following sequential steps: data preprocessing (Section III-B), **fine-grained relational graph construction** (Section III-C), **graph-conditioned replay** (Section III-D), and **representation alignment** (Section III-E). For data preprocessing, each client standardizes its local MRI and DCE-MRI scans, segments disease-specific regions, and extracts spatial morphology and temporal enhancement descriptors. For fine-grained relational graph construction, each client extracts privacy-safe disease summaries s_i^k (spatial relations) and r_i^k (temporal relations) from local annotations and sends them to the server. The server constructs a fine-grained relational graph

$$G^k = \Phi\big(\{s_i^k\}_{i=1}^N; \{r_i^k\}_{i=1}^N\big), \qquad g_i^k = G_{i,:}^k,$$

where G_{ij}^k measures the fine-grained relation between clients

C_i and C_j, and g_i^k denotes the graph row assigned to client C_i. Section III-C defines Φ using fine-grained spatial and temporal attention. For graph-conditioned replay, each client maintains an encoder-generator-predictor triplet (E_i, G_i, F_i). The encoder extracts graph-conditioned representations from current-task samples, while the generator synthesizes representations for current and previous tasks conditioned on labels and graph rows. The predictor learns jointly from real and replayed representations to adapt to the current task while preserving prior knowledge. For representation alignment, a global discriminator D reconstructs graph rows from noise-perturbed real and synthetic representations. The encoders and generators are trained adversarially against D to reduce graph-specific biases and align representations across clients and tasks. Overall, *BreastG-FCL* uses fine-grained relational graphs to guide replay and alignment under evolving *TCGA-BRCA* client distributions without sharing raw patient data.

B. Data Preprocessing

We preprocess breast MRI and DCE-MRI data x from Task k of client C_i using bias-field correction, intensity normalization, spatial resampling, and inter-phase registration, yielding the standardized scan $I_{i,x}^k$. We segment four disease-relevant regions, $\{M_{i,x,\omega}^k\}_{\omega\in\Omega_{\text{sp}}}$, where $\Omega_{\text{sp}} = \{\text{core}, \text{rim}, \text{peri}, \text{bpe}\}$, and aggregate their morphology and enhancement features into the client-level spatial representation s_i^k. For DCE-MRI, we normalize the temporal signal to obtain the enhancement curve $E_{i,x}^k$ and extract $r_{i,x}^k = [\text{wash-in}\|\text{wash-out}\|\text{TTP}\|\text{SER}]$. We aggregate these features into the client-level temporal representation r_i^k. Thus, each client-task pair is represented by the processed spatial and temporal data (s_i^k, r_i^k).

C. Fine-grained Graph Generation

For each Task k, client C_i summarizes its local breast cancer annotations into a spatial morphology vector s_i^k and a temporal kinetics vector r_i^k. We model these two signals separately because tumor morphology characterizes the current disease phenotype, whereas DCE-MRI kinetics describe how contrast-enhancement patterns evolve across tasks. The server first computes fine-grained spatial attention as

$$\alpha_{ij}^k = \text{softmax}_j\left(a_{\theta_s}(s_i^k, s_j^k)\right),$$

where the attention network, $a_{\theta_s}(\cdot)$, measures the similarity between the tumor morphology summaries of clients C_i and C_j. The softmax operation normalizes these similarities across candidate clients, allowing α_{ij}^k to represent the relative spatial relevance of C_j to C_i.

To capture temporal evolution rather than relying only on the current task, we concatenate the kinetic summaries within a sliding window of length m: $R_i^k = \left[r_i^{\max(1,k-m+1)}\|\cdots\|r_i^k\right]$. The sliding window preserves recent enhancement dynamics while limiting the influence of outdated tasks and controlling the representation size. The temporal attention is computed as

$$\beta_{ij}^k = \text{softmax}_j\left(b_{\theta_t}(R_i^k, R_j^k, m)\right),$$

where the attention network, $b_{\theta_t}(\cdot)$, measures the similarity between the recent DCE-MRI enhancement trajectories of clients C_i and C_j.

Finally, we combine the two attention scores through multiplicative fusion:

$$G_{ij}^k = \frac{\alpha_{ij}^k \beta_{ij}^k}{\sum_{\ell=1}^N \alpha_{i\ell}^k \beta_{i\ell}^k + \varepsilon_0}, \qquad g_i^k = G_{i,:}^k,$$

where $\varepsilon_0 > 0$ ensures numerical stability. Multiplicative fusion assigns a large edge weight only when two clients are similar in both tumor morphology and temporal enhancement dynamics, thereby suppressing relations supported by only one type of evidence. Thus, a larger G_{ij}^k indicates that clients C_i and C_j exhibit jointly consistent breast tumor morphology and DCE-MRI kinetic patterns.

D. Graph-Conditioned Replay

After graph generation, the graph row serves as a client-specific relational context, summarizing how C_i is related to other clients in terms of tumor morphology and DCE-MRI enhancement dynamics. Conditioning local learning on this context helps account for cross-client heterogeneity rather than treating all clients as having the same data distribution. Each client maintains an encoder-generator-predictor triplet (E_i, G_i, F_i). For real samples from the current task, the encoder produces graph-conditioned latent representations:

$$\mathcal{E}_{i,\text{real}}^k = E_i(x, g_i^k), \qquad (x, y) \in \mathcal{D}_i^k.$$

Conditioning the encoder on g_i^k encourages the extracted representations to capture both sample-level information and the current relational structure of client C_i.

For replay, the generator produces synthetic latent representations conditioned on random noise, label information, and historical graph rows:

$$\mathcal{E}_{i,\text{syn}}^j = \mathcal{G}_i(z, y, g_i^j), \qquad z \sim \mathcal{N}(0, I), \quad j \le k.$$

Here, z introduces sample diversity, y preserves class semantics, and g_i^j retains the client- and task-specific relational context of Task j. Thus, the generator replays relation-aware representations instead of generic class-conditioned representations.

The predictor F_i is trained on real representations from the current task and synthetic representations from the current and previous tasks. The real representations support adaptation to newly observed clinical distributions, whereas the synthetic representations preserve previously acquired knowledge without storing raw historical data. In this way, graph-conditioned replay balances continual adaptation and knowledge retention under heterogeneous *TCGA-BRCA* client distributions.

E. Representation Alignment

BreastG-FCL adopts a server-side global discriminator D to align real and synthetic latent representations. Since each local model observes only its own data distribution, the global discriminator provides a shared alignment objective across clients and tasks without requiring access to raw patient data. Unlike a standard discriminator that predicts

only real/fake labels, D predicts the fine-grained graph row associated with each representation. The graph row provides richer supervision by characterizing client-specific relations in tumor morphology and DCE-MRI kinetics, thereby allowing D to identify relational biases.

The clients upload noise-perturbed real and synthetic representations together with their corresponding noise-perturbed graph rows:

$$\hat{\mathcal{E}}^k_{i,\text{real}} = E_i(x, g^k_i) + \tau_i, \qquad \hat{g}^k_i = g^k_i + \tau_g,$$
$$\hat{\mathcal{E}}^j_{i,\text{syn}} = \mathcal{G}_i(z, y, g^j_i) + \tau_i, \qquad \hat{g}^j_i = g^j_i + \tau_g, \quad j \le k,$$

where τ_i and τ_g denote privacy-preserving perturbations applied to the latent representations and graph rows, respectively. These perturbations reduce the exposure of client-specific information during server-side alignment.

The discriminator is trained to reconstruct $\hat{g}^k_i$ and $\hat{g}^j_i$ from the corresponding uploaded representations, making it sensitive to client- and task-specific relational patterns. In contrast, the local encoders and generators are trained adversarially to make this reconstruction difficult. Applying the same alignment objective to both real and synthetic representations further encourages them to occupy a consistent latent space. Together with the supervised prediction objective, this design preserves breast-cancer-discriminative information while improving representation alignment across heterogeneous clients and tasks.

IV. Evaluation

We evaluate *BreastG-FCL* on breast cancer radiogenomic modeling under the federated continual learning setting. Based on the evaluation protocol of previous studies [10], our experiments aim to answer three questions: (1) whether *BreastG-FCL* improves prediction performance over existing FCL baselines; (2) whether disease-aware spatial and temporal attention improves graph-conditioned replay; and (3) whether the learned graph captures clinical information. This work is implemented with NVFlare [19] and will be publicly available.

A. Experimental Settings

Dataset. We use *TCGA-BRCA* [15] as the main healthcare dataset, which contains 1098 patient cases. Each patient case contains high-dimensional gene expression features, clinical covariates, subtype labels, and demographic information. Under the current GDC STAR-Counts processing based on GENCODE v36, each gene-expression profile contains 60,660 gene-level measurements, including 19,962 protein-coding genes. We obtain the accompanying radiological data from the *TCGA-BRCA* collection in TCIA [20].

FCL Setting. Since *TCGA-BRCA* is not naturally federated, we partition samples into N clients according to clinically meaningful non-IID factors, such as age. Each client observes a sequence of K tasks. Tasks are constructed to simulate evolving healthcare distributions, such as subtype shift, stage shift, imaging-kinetics shift, or mixed clinical non-IID shift.

Baselines. We compare *BreastG-FCL* with representative FCL methods, including FedAvg-Adv, FedCIL [17], MFCL [21], AF-FCL [10], C^{2}Prompt [22], and GFedCL [12].

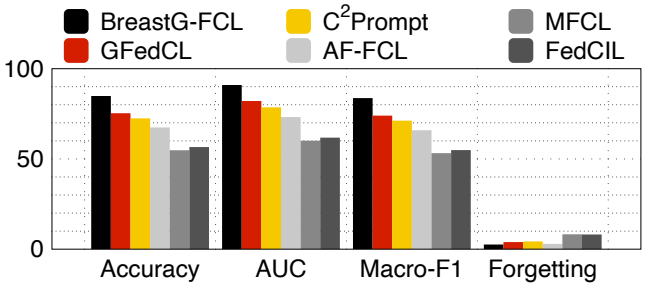

Fig. 4: Results on *TCGA-BRCA*. Higher Accuracy, AUC, and Macro-F1 are better; lower Forgetting is better.

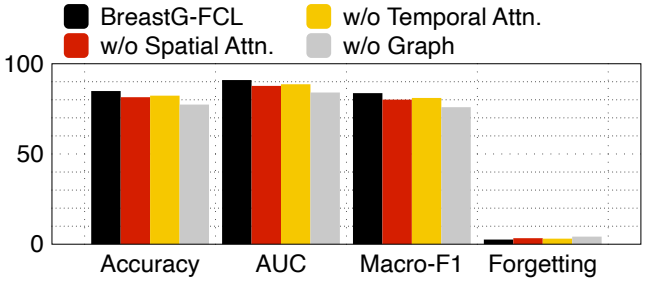

Fig. 5: Ablation study on *TCGA-BRCA*. "Attn." stands for "Attention."

Metrics. For classification, we report Accuracy, AUC, Macro-F1, and Forgetting. *Accuracy*, *AUC*, and *Macro-F1* measure downstream prediction performance, while *forgetting* measures performance degradation on previous tasks:

$$\text{Forgetting} = \frac{1}{K-1}\sum_{j=1}^{K-1}\left(\max_{l\in\{j,\ldots,K-1\}} A_{l,j} - A_{K,j}\right),$$

where $A_{k,j}$ denotes the average performance on task j after training up to task k.

B. Main Results on TCGA-BRCA

As Figure 4 shows, *BreastG-FCL* achieves the best results across all metrics, with 84.87% accuracy, 90.96% AUC, 83.71% Macro-F1, and only 2.58% forgetting. Compared with GFedCL, it improves accuracy, AUC, and Macro-F1 by 9.55, 8.86, and 9.71 percentage points, respectively, while reducing forgetting by 1.35 percentage points. These gains show that generic relations derived from client updates cannot fully characterize breast cancer heterogeneity. Conditioning replay on tumor morphology and DCE-MRI kinetics provides more clinically meaningful relations, thereby improving both continual adaptation and prior-task knowledge preservation.

C. Ablation Study

Figure 5 shows that all components contribute to the final performance. Removing clinical spatial attention reduces accuracy by 3.41 percentage points results in a 2.56-percentage-point decrease This indicates that tumor morphology provides the primary relational signal, while DCE-MRI kinetics supplies complementary dynamic information. Removing the entire

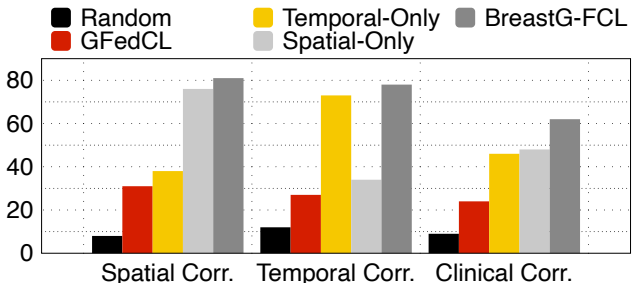

Fig. 6: Correlation between learned graph weights and clinical similarities. "Corr." stands for "Correlation."

graph produces the largest degradation, reducing accuracy to 77.38% and increasing forgetting to 4.21%. These results demonstrate that disease-aware graph conditioning is essential for generating replay representations that preserve previous knowledge while adapting to heterogeneous clients.

D. Clinical Graph Analysis

To verify whether the learned graph captures clinically meaningful relations, we compare graph edge weights with client-level clinical similarities. For each task k, we compute the Pearson correlation between graph weights $\{G_{ij}^k\}_{i\neq j}$ and three similarity scores: $S_{ij}^{k,\text{spatial}} = \exp(-L_2(s_i^k, s_j^k))$, $S_{ij}^{k,\text{temporal}} = \exp(-L_2(R_i^k, R_j^k))$, $S_{ij}^{k,\text{clinical}} = \exp(-L_2(c_i^k, c_j^k))$, where s_i^k denotes the morphology summary, R_i^k denotes the DCE-MRI kinetics history, and c_i^k denotes clinical composition. A higher positive correlation indicates that *BreastG-FCL* assigns larger edge weights to clinically similar clients.

As shown in Figure 6, the random graph has limited correlation with all clinical similarities, while the original GFedCL graph exhibits only moderate correlations. The spatial-only and temporal-only graphs achieve high correlations primarily within their respective dimensions, confirming that morphology and DCE-MRI kinetics capture complementary relationships. By integrating both signals, the *BreastG-FCL* graph achieves the highest spatial, temporal, and clinical correlations of 0.81, 0.78, and 0.62, respectively. These results indicate that the learned graph connects clients with similar phenotypes and also reflects broader molecular and clinical cohort structures.

V. Conclusion

We proposed *BreastG-FCL*, a disease-aware federated continual learning framework for breast cancer radiogenomic modeling. By constructing relational graphs from tumor morphology and DCE-MRI enhancement kinetics, *BreastG-FCL* improves predictive performance while mitigating catastrophic forgetting and cross-client heterogeneity. The framework may also be adaptable to other contrast-enhanced radiogenomic applications, such as prostate cancer.

References

[1] Z. Liu *et al.*, "Radiogenomics: A key component of precision cancer medicine," *British Journal of Cancer*, 2023.

[2] A. Saha *et al.*, "A machine learning approach to radiogenomics of breast cancer: a study of 922 subjects and 529 dce-mri features," *British Journal of Cancer*, 2018.

[3] J. Yoon *et al.*, "Federated continual learning with weighted inter-client transfer," in *ICML*, 2021.

[4] S. Wu *et al.*, "DCE-MRI background parenchymal enhancement quantified from an early versus delayed post-contrast sequence: Association with breast cancer presence," *Scientific Reports*, 2017.

[5] K. He *et al.*, "Deep residual learning for image recognition," in *CVPR*, 2016.

[6] United States Congress, "Health insurance portability and accountability act of 1996," Public Law 104-191, 1996, 110 Stat. 1936. [Online]. Available: https://govinfo.gov

[7] H. B. McMahan *et al.*, "Communication-efficient learning of deep networks from decentralized data," in *AISTATS*, 2017.

[8] M. De Lange *et al.*, "A continual learning survey: Defying forgetting in classification tasks," *IEEE TPAMI*, 2022.

[9] J. Dong *et al.*, "Federated class-incremental learning," in *CVPR*, 2022.

[10] W. Abudukelimu *et al.*, "Accurate forgetting for heterogeneous federated continual learning," in *ICLR*, 2024.

[11] R. He *et al.*, "AFL: A single-round analytic approach for federated learning with pre-trained models," in *CVPR*, 2025.

[12] Q. Yu *et al.*, "Gfedcl: Graph-based federated continual learning with spatial and temporal awareness," in *ICML*, 2026.

[13] J. Zhou *et al.*, "Weakly supervised 3d deep learning for breast cancer classification and localization of the lesions in mr images," *JMRI*, 2019.

[14] Y. Zhu *et al.*, "Deciphering genomic underpinnings of quantitative MRI-based radiomic phenotypes of invasive breast carcinoma," *Scientific Reports*, 2015.

[15] The Cancer Genome Atlas Network, "Comprehensive molecular portraits of human breast tumours," *Nature*, 2012.

[16] H. Li *et al.*, "Quantitative MRI radiomics in the prediction of molecular classifications of breast cancer subtypes in the TCGA/TCIA data set," *npj Breast Cancer*, 2016.

[17] D. Qi *et al.*, "Better generative replay for continual federated learning," in *ICLR*, 2023.

[18] N. Tishby *et al.*, "The information bottleneck method," in *AACCCC*, 1999.

[19] H. R. Roth *et al.*, "Nvidia flare: Federated learning from simulation to real-world," 2022.

[20] K. Clark *et al.*, "The cancer imaging archive (tcia): Maintaining and operating a public information repository," *Journal of Digital Imaging*, 2013.

[21] B. Sara *et al.*, "A data-free approach to mitigate catastrophic forgetting in federated class incremental learning for vision tasks," in *NeurIPS*, 2023.

[22] K. Xu *et al.*, "c^2 prompt: Class-aware client knowledge interaction for federated continual learning," in *NeurIPS*, 2025.