Mesopotamian journal of Big Data
Vol. **(2025)**, 2025, **pp**. 222–240
DOI: https://doi.org/10.58496/MJBD/2025/015 , ISSN: 2958-6453
https://mesopotamian.press/journals/index.php/BigData

Research Article

Shared Generator-Based Serverless Multimodal Federated Learning For Medical Image Analysis

Mohammed Adel Al-Shaher[1], Ahmed Saihood[1,*], Mustafa Asaad Hasan[1]

[1] *College of Computer Science and Mathematics, University of Thi-Qar, ThiQar, Iraq*

ARTICLE INFO

Article History

Received 25 Jun 2025
Revised 18 Jul 2025
Accepted 27 Aug 2025
Published 22 Sep 2025

Keywords

FL
SGS-FL
Medical imaging
Lung nodules
Multimodality models

ABSTRACT

Medical image analysis constitutes the foundation of the diagnosis of potential life-threatening conditions such as lung cancer. Nevertheless, AI model construction in this context is hindered by strict privacy regulations (e.g., HIPAA, GDPR), the variability of imaging protocols, and the paucity of large annotated datasets. These barriers constrain centralised machine learning and dampen interdisciplinary research. To overcome such barriers, this paper introduces shared generator-serverless federated learning (SGS-FL), a decentralised multimodal medical imaging framework. By employing a shared generator and multidiscriminator architecture, SGS-FL eliminates centralised dependency via cross-modal synthesis, while the communication burden is reduced by embedding a sharing protocol. By employing latent space aggregation with attention and independent component analysis, the interpretability, fairness, and relevance of features are improved. Experimental evaluation was conducted across three lung cancer datasets: LIDC-IDRI CT scans (≈1,018 cases), NODE21 chest radiographs (~10,000 images), and NSCLC radiogenomic PET-CT images (~211 patients). By employing 10-fold cross-validation with 10 independent iterations, SGS-FL achieved 92.5% ± 1.2 accuracy, 0.83 ± 0.02 Dice coefficient, 0.946 ± 0.01 area under the curve, and 21.5 ± 1.1 Frechet inception distance (FID), significantly surpassing benchmark state-of-the-art schemes such as FedACS (~88%) and Federated Transfer Learning (~89%) ($p < 0.01$). The results indicate that SGS-FL achieves superior scalability, interpretability, and performance and constitutes a sound paradigm of privacy-friendly and clinically trustworthy AI in medical imaging.

1. INTRODUCTION

Medical image analysis forms the cornerstone of modern diagnostic workflows, offering critical insights into diseases such as lung cancer, cardiovascular conditions, and neurodegenerative disorders. However, developing AI for medical imaging presents daunting challenges. These range from stringent data protection regulations, such as the HIPAA and GDPR, data islands within institutions, and wide variability in imaging modalities and annotation standards. As a result, institutions are reluctant or incapable of sharing raw medical data, limiting the possibility of centralising machine learning models on voluminous, heterogeneous datasets. Therefore, collaborative but privacy-preserving machine learning architectures are needed more than ever before in healthcare AI [1].

Federated learning (FL) [2] offers an ideal solution, accommodating model training while not exchanging raw data. While promising, most existing FL techniques are not sufficient when applied to realistic medical imaging settings. They often employ a central coordinating server, causing a single point of failure and administrative chokepoints. Most techniques also assume homogeneous data distributions, an unrealistic assumption given the heterogeneity of medical imaging protocols and patients across institutions. For these reasons, the existing federated techniques are plagued with poor model performance, limited generalizability to unseen data, and limited interpretability—a necessary requirement on clinical deployment fronts where the trustworthiness of AI systems matters most [3]–[5].

FL methods such as Federated Adaptive Client Sampling (FedACS) [5], Federated Transfer Learning (FTL) [4], and homomorphic encryption-based FL [3] provide some interesting innovations but nevertheless suffer from several key weaknesses. FedACS imposes drastic computational and communication overheads as a side effect of dynamic strategies for client selection. FTL demands significant energy resources and fine-tuning and hence is not realistic for organisations with weak computational ability. Homomorphic encryption methods, despite their strong guarantees of privacy, inflict drastic delays as a side effect of the encryption and decryption processes. Furthermore, most of these methods do not provide transparency of decisions, with no special facilities for visualising and verifying features affecting AI predictions [6]–[8].

*Corresponding author. Email: ahmed.alisiehood@utq.edu.iq

The Shared Generator Serverless Federated Learning (SGS-FL) system, as formulated in this article, addresses these main gaps with an entirely decentralised, serverless approach to multiparty medical imaging AI. SGS-FL avoids central coordination requirements with peer-to-peer model updates. SGS-FL uses a shared generator model on all clients, with discriminators on each node to enable modality-specific learning. This architecture accommodates the production of cross-modal synthetic data supporting the learning processes while ensuring confidentiality. Embedding-sharing communications reduce bandwidth requirements further with short, low-dimensional vector message sharing, as a substitute for sharing full model parameters, and offer efficiency and scalability despite modest resources.

The motivation behind SGS-FL is to overcome the bounds of current federated learning protocols. Most of the current methods fall short of achieving a balance between the trade-offs between privacy, effectiveness of communication, generalizability, and explainability [7], [9], [10]. SGS-FL caters to delivering a one-stop solution with elevated accuracy over divergent data, visual explanations of decisions with attention-based modalities, and compatibility with low communication overhead. With equity and balanced integration of client submissions, SGS-FL ensures that all nodes participating—be their size of data large or small or their modality modal or multimodal—play an influential role during model updation, overcoming fairness issues inherent to prior techniques [6], [7], [11], [12].

SGS-FL achieves these goals with an architectural and algorithmic combination. A typical generator produces realistic cross-modal data with embedding vectors of diverse client modalities, and a multidiscriminator setup per node ensures that synthetic data are rigorously evaluated for realism and pertinence. Latent space consolidation with attention identifies and prioritises the most information-rich features among clients and decentralises feature selection with independent component analysis (ICA) [13] to adjust model updates. These components work towards developing a model that is not only accurate but also explainable, generalisable, and efficient [14], [15].

Medical image analysis currently underlies mainstream diagnostic pipelines, facilitating early identification and control of severe diseases such as lung carcinoma, cardiovascular diseases, and neurodegenerative diseases. However, creating trustworthy artificial intelligence (AI) systems in medical imaging is still challenging because of data privacy statutes (e.g., HIPAA, GDPR) and local data silos and because of variations in imaging modalities and annotation practices. These issues significantly hinder the practicality of centralised machine learning in large, heterogeneous datasets.

Recent designs such as AdFed (Decentralised FL) and FACL (Federated Attention Consistent Learning) attempt to handle decentralisation and interpretability but are limited. AdFed has difficulty dealing with high-order regularisation and poor accuracy under heterogeneous data, and FACL guarantees attention consistency but involves high complexity in large-scale implementations. In contrast, SGS-FL provides a shared generator with multiple discriminator nodes and embedding-sharing communication and allows effective, interpretable, and nondiscriminative multimodal learning with no central coordination.

The current FL methods are subject to several issues: (1) Reliance on central servers (a single point of failure, administrative bottlenecks). (2) Unrealistic assumption of homogeneous data distributions. (3) High computational and communication overhead (e.g., FedACS). (4) Excessive energy costs (e.g., federated transfer learning (FTL)). (5) Slow training from encryption-based methods. (6) Lack of interpretability and transparency of prediction. (7) Unlike earlier decentralised approaches such as AdFed, which remove the central server but lack accuracy under heterogeneous datasets, or attention-centric approaches such as FACL, which enhance interpretability but are highly complex and have poor scalability, the proposed SGS-FL framework achieves a balance among decentralisation, fairness, and efficiency. By incorporating a shared generator, a multidiscriminator structure, and embedding-sharing communication along with attention-controlled feature aggregation, SGS-FL overcomes the primary limitations of current federated learning approaches. This placement reveals the innovation of our approach, with superior robustness, interpretability, and scalability for multimodal medical image analysis.

The proposed SGS-FL offers a novel federated learning paradigm tailored for medical imaging applications, where privacy, transparency, and adaptability are essential. The design of SGS-FL supports collaborative training across diverse institutions and modalities without sacrificing performance or operational feasibility. By addressing the core shortcomings of prior FL methods, SGS-FL represents a significant step forward in enabling scalable, explainable AI for healthcare [16]. The main contributions of this article are as follows:

- We propose SGS-FL, a fully serverless and decentralised federated learning framework that integrates multimodal medical data while preserving privacy.
- We introduce a shared generator with a multidiscriminator architecture that facilitates cross-modal synthesis and ensures robust generalisation.
- We design an embedding-sharing communication protocol that significantly reduces bandwidth consumption compared with weight-sharing methods.

- We incorporate attention-based latent space fusion and decentralised feature selection to improve interpretability, fairness, and feature relevance across clients.
- We provide comprehensive experimental validation on lung cancer datasets (CT, PET, and X-ray), which show superior performance in classification, segmentation, and synthetic data quality relative to state-of-the-art federated learning baselines.

The remainder of this paper is structured as follows: Section 2 reviews related work in federated learning for medical imaging; Section 3 presents the proposed SGS-FL methodology; Section 4 describes the experimental setup and results; Section 5 discusses findings, limitations, and implications; and Section 6 concludes the paper.

2. RELATED WORKS

FL has gained significant attention as a privacy-preserving alternative to centralised machine learning in medical image analysis. One early approach, FedACS [5], aimed to improve resource utilisation and model performance by dynamically selecting clients during training. Although effective in balancing the computational load, FedACS introduced considerable overhead because of its complex client sampling procedures. Similarly, [4] targeted medical image classification with a focus on privacy preservation. While FTL has shown promise in terms of classification accuracy, it comes at the cost of high energy consumption and requires extensive fine-tuning to manage heterogeneous data distributions.

Several methods have sought to extend FL into the domain of medical image segmentation. Federated semisupervised learning with pseudolabel denoising [8] and mixed supervised federated learning [17] attempt to address limited labelled data by incorporating pseudolabelling strategies. Although these approaches improved segmentation, their reliance on pseudolabels introduced sensitivity to label noise, and the implementations were often complex. FKD [12] provides an alternative by sharing distilled knowledge rather than full models or data, improving communication efficiency but sometimes sacrificing critical information necessary for fine-grained segmentation tasks [18].

An overview of available FL techniques presents both promising developments and persisting shortcomings. Methods such as Federated Adaptive Client Sampling (FedACS) [5] and Federated Transfer Learning (FTL) [4] seek better efficiency and improved resource use. Nevertheless, their computational overhead and energy costs make them inapplicable in numerous limited-resource medical centers. Similarly, semisupervised FL with pseudolabel denoising [8] and mixed supervised FL [20] extend the usability of small annotated sets, but pseudolabel dependences add noise sensitivity, causing poor-quality segmentation. Privacy-oriented architectures, such as homomorphic encryption-based FL [3] and blockchain-based FL with causal learning [19], ensure data integrity at the price of high latency, encryption overhead, and system complexity costs. Without exception, these approaches reflect privacy guarantees, computational efficacy, and real-time usability compromises, which are still significant barriers to clinical applicability.

There have also been architectural advances, such as multibranch prototype FL [21] and federated fusion learning with attention [9], which enhance non-IID and heterogeneous data handling but have higher architectural complexity and communication costs, hindering scalability. Federated knowledge distillation (FKD) [12] improved the efficiency of distilled model sharing but was associated with the risk of losing relevant fine-grained information. In interpretability, Federated Attention Consistent Learning (FACL) [7] has advanced attention reliability in prostate cancer research, whereas Decentralised FL (AdFed) [10] has removed enabling central servers in the quest for enhanced resilience. However, FACL has high complexity in large-scale deployments, and AdFed has accuracy problems under heterogeneous datasets with intricate regularisation. More recently, LWR-Net [30] introduced scalable multitask adaptation with nonretraining, providing efficiency and domain-agnostic generalizability, although it was not explicitly devised with federated settings in mind. In general, these works highlight the need for a solution that balances privacy, interpretability, computational efficiency, and fairness across multimodal, heterogeneous sets, and SGS-FL is introduced here to fulfil that need [19].

Privacy-preserving FL methods such as horizontal federated learning with homomorphic encryption [20], [21] and blockchain-enhanced Federated learning with causal learning [22][23] emphasise data integrity and security. While these frameworks achieve strong privacy protection, they face significant drawbacks in practice. Homomorphic encryption methods slow training because of encryption overhead, and blockchain-based solutions introduce latency and additional system complexity. Moreover, ensuring fairness and managing large-scale deployments remains challenging, as noted in comparative studies [24] and evaluations for specific tasks such as breast density classification [25].

Efforts to enhance generalizability and manage data heterogeneity include multibranch prototype federated learning (FEDMBP) [26] and federated fusion learning with attention mechanisms [9]. These methods address the challenge of non-IID data distributions by introducing architectural innovations such as multibranch models and attention fusion. However, these enhancements often come at the cost of increased model complexity and greater communication needs. Similarly, FACL [5] focused on consistent attention across nodes for prostate cancer diagnosis, but managing large-scale attention consistency proved difficult, and performance gains were limited in highly imbalanced or heterogeneous datasets [27].

Moreover, many studies have explored FL in specific medical domains (see Table 1), including CNN-based Federated Learning for Cervical Cancer Classification [28], Federated Active Learning for Skin-Lesion Classification [29], and FedEYE for Ophthalmology [30]. These platforms demonstrated the feasibility of applying FL across various specialties but highlighted challenges related to scalability, flexibility, and periodic annotation overhead. AdFed [10] presented a step toward fully decentralised architectures, yet it faced complexities in regularisation and privacy management. Collectively, these related works underscore the need for an FL framework that not only preserves privacy but also effectively handles multimodal data, reduces communication costs, ensures fairness, and offers transparent, interpretable outputs—gaps that the proposed SGS-FL framework is designed to address.

Table 1: Summary of existing federated learning methods applied to medical imaging, their applications, evaluation metrics, and key limitations.

Ref.	Method of Federated Learning	Application	Evaluation Metrics	Issues and Shortcomings
[5]	Federated Learning with Adaptive Client Sampling	Medical Image Analysis, Classification	Accuracy, training time, resource utilisation	Requires careful selection of clients and computational overhead.
[4]	Federated Transfer Learning	Medical Image Classification, Privacy-preserving analysis	Classification accuracy, energy consumption, and privacy levels	Energy consumption and privacy levels might still need optimisation.
[8]	Federated Semi-Supervised Learning with Pseudo-Label Denoising	Medical Image Segmentation	Dice coefficient, IoU	The quality of pseudolabels can affect performance, complex implementation.
[31]	Comparative study of different Federated Learning Techniques	Secure and privacy-preserving machine learning on medical datasets	Model accuracy, data privacy levels, computational cost, robustness	Comparison complexity and ensuring fairness in evaluation.
[32]	Federated Learning with Causal Learning and Blockchain	Explainable Medical Image Analysis	Performance, explainability, data integrity and security	Blockchain integration may introduce latency; complexity in causal learning.
[3]	Horizontal Federated Learning with Homomorphic Encryption	Lung Image Classification	Classification accuracy, encryption overhead, and privacy protection levels	Encryption overhead can slow down training, complexity in implementation.
[33]	Mixed Supervised Federated Learning	Medical Image Segmentation	Dice coefficient, IoU	Handling mixed supervision can be challenging, as pseudolabel accuracy.
[26]	Multi-Branch Prototype Federated Learning	Medical Image Processing	Accuracy, data heterogeneity management, model generalisation	Managing multiple branches and prototypes adds complexity.
[12]	Federated Learning with Knowledge Distillation	Medical Image Segmentation	Segmentation accuracy, communication efficiency, and privacy metrics	Knowledge distillation may lead to a loss of some information.
[34]	Generative Adversarial Networks (GANs)	Medical Image Synthesis, Enhancement, Augmentation, and Segmentation	Image quality, realism, and enhancement capabilities	GANs can be challenging to train due; stability issues.
[35]	Federated Active Learning with ensemble-entropy-based AL	Skin-lesion classification	Performance on dermoscopic datasets	Depends heavily on the annotation strategy; periodic execution overhead.
[36]	Model-level Attention and Batch-Instance Style Normalisation	Medical Image Segmentation	Dice similarity coefficient	Requires balancing attention mechanisms and style normalisation.
[37]	Multimodal Federated Learning	Medical image reporting	Accuracy and quality of medical image reports	Combining multimodal data increases complexity; data heterogeneity.
[38]	Federated Learning evaluation for breast density classification	Breast density classification	Linear kappa score on test and external datasets	Ensuring fair evaluation across diverse datasets.
[9]	Federated Fusion Learning with Attention Mechanism	Medical Image Analysis	Generalizability, performance on MedMNIST datasets	Handling non-IID data and statistical heterogeneity.
[39]	Federated Transfer Learning	Kidney disorder identification	Precision, efficacy of renal abnormality detection	Transfer learning models need fine-tuning, data heterogeneity.
[7]	Federated Attention Consistent Learning (FACL)	Prostate cancer diagnosis and Gleason grading	AUC, Kappa score	Ensuring attention consistency, managing large-scale data.
[40]	CNN-based Federated Learning	Cervical cancer classification	Test accuracy in IID and non-IID settings	IID vs non-IID performance disparity; small dataset issues.
[10]	Decentralised Federated Learning (AdFed)	Cancer survival prediction	AUC on survival prediction datasets	Regularisation and ensuring privacy protection can be complex.
[30]	Federated Learning platform for ophthalmic disease image classification	Ophthalmology	Applicability on various neural networks and classification tasks	Deployment specifics might limit scalability and flexibility.
[41]	LWR-Net (Learning Without Retraining)	Multitask medical image analysis, general deep learning applications	Accuracy, generalisation across domains, computational efficiency	Eliminates need for retraining when adapting to new tasks; efficient feature fusion; supports domain-agnostic generalisation.

3. PROPOSED SHARED GENERATOR-BASED SERVERLESS FL (SGS-FL)

The proposed method addresses the following issues: (1) Almost all methods suffer from the complexity of combining multimodal data and data heterogeneity. The proposed method effectively handles data heterogeneity through robust aggregation and the attention-based approach in multiclient-multimodality fusion. (2) Client feature selection may

lead to significant feature loss through latent space collection; this method automatically weighs the importance of each client's contribution, reducing the need for explicit client selection and computational overhead. (3) Almost all the proposed methods face problems in ensuring fairness in evaluating clients; in this article, the attention-based approach provides a unified framework for assessing contributions from different clients, ensuring fairness. (4) The complexity of pseudolabels can affect performance; this attentive aggregation ensures that the best latent spaces are used, mitigating the impact of poor-quality pseudolabels. (5) Most methods depend heavily on annotation strategies and periodic execution overhead. In this work, the attention-based approach reduces the dependency on annotation strategies by focusing on the most informative features.

This section explains the proposed SGS-FL (see Fig.1), a fully serverless, generalisable, decentralised multimodal learning framework. It shares a single generator model across all clients, while each client maintains multiple unique discriminators. Each client has N discriminators (including one for its data), where N is the number of clients (nodes). No centralised server is involved. Instead, model updates are collaboratively exchanged and averaged directly between clients.

SGS-FL is motivated by two innovations: 1) Generator sharing with a multidiscriminator architecture, which enables cross-modality synthesis and distributed training without central coordination. 2) Embedding-based communication, where low-dimensional embeddings are exchanged between clients instead of full model parameters, significantly reduces the communication cost and enhances privacy. Each client operates with a shared generator and multiple discriminators, one per participating node (including itself), enabling modality specialisation and generalisation across the system.

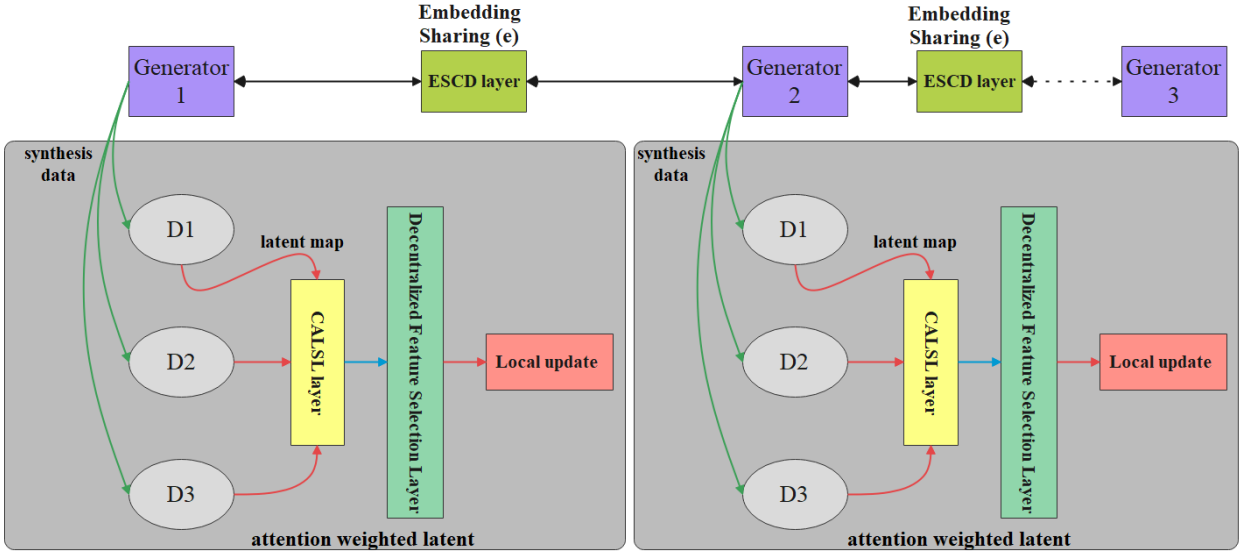

Fig.1:. The SGS-FL architecture shows the shared generator and multidiscriminator design.

3.1. Generator Sharing and Multidiscriminator Setup

This process also reduces the communication overhead because clients need only transmit low-dimensional embeddings and not all model parameters or original data. Therefore, the system is efficient and capable of maintaining privacy. The adoption of several discriminators also ensures that every client can check samples it produces against all the modalities used, hence increasing resistance against data heterogeneity and guaranteeing equal participation of different clients. These design choices collectively provide the foundation for facilitating a serverless and scalable federated learning system.
Considering that we have N clients, each client $i \in \{1, \ldots, N\}$ maintains the shared generator's identical architecture across all the clients. A set of N discriminators $\{D_{i,j}\}_{j=1}^{N}$ at each node. The generator G takes as input a concatenated vector of a random latent code $z \sim N(0,1)$ and an embedding vector $e_j \in R^d$ corresponding to client j. This will produce the synthetic image shown in Eq. (1).

$$\tilde{x}_{i,j} = G([z; e_j]) \ldots . (1)$$

This allows any client to generate data resembling another client's modality by using its embedding.

3.2. Discriminator Loss and Feedback Latent Space

Each discriminator $D_{i,j}$ is trained to distinguish between real images $x_j \sim X_j$ from client j's modality and fake images $\tilde{x}_{i,j}$ generated by G. The binary cross-entropy loss for the discriminator $D_{i,j}$ is defined in Eq. (2).

$$\mathcal{L}_{D_{i,j}} = -E_{x_j \sim X_j}[log\ D_{i,j}(x_j)] - E_{z\sim N(0,1)}[log(1 - D_{i,j}(\mathrm{G}([z; e_j])))] \ldots\ldots (2)$$

Each $D_{i,j}$ also outputs a latent feature map $Z_{i,j} \in R^{h\times w}$ from its penultimate layer. All such maps at client i are stacked to form Z_i **via** Eq. (3)**.**

$$Z_i = [Z_{i,1}, Z_{i,2}, \ldots . Z_{i,N}] \in R^{h\times w\times N} \ldots . . (3)$$

This stacked latent space encodes feedback from each peer-discriminator, carrying modality-specific encoding.

3.3. Embedding a sharing-based collaborative discriminator (ESCD)

Clients exchange only their embedding vectors e_j with each other, not the full model weights. This is facilitated by a learnable embedding table where each $E = \{e_1, \ldots . . , e_N\}\ e_j \in R^d$ represents one client or modality.

The embedding-sharing mechanism plays a vital role in enabling collaboration across clients while preserving privacy and minimising communication overhead. By exchanging only low-dimensional embedding vectors rather than full model weights or raw data, each client gains the ability to generate synthetic samples that reflect the distribution of other modalities present in the federation. The embeddings $e_j \in R^d$ act as compact and abstract representations of the underlying modality-specific knowledge of client j. When combined with latent codes $z \backsim N\ (0,1)$, these embeddings guide the shared generator G to produce synthetic data that align with the characteristics of the target modality. This setup allows any client i to synthesise data resembling the modality of any other client j, thereby facilitating rich cross-modal learning.

The collaborative discriminators at each node are specifically designed to leverage these embeddings for modality-specialised supervision (see *Fig. 2*). Every discriminator $D_{i,j}$ is tasked with distinguishing between real images $x_{i,j}$ from client j's modality and the synthetic images $\hat{x}_{i,j} = G([z; e_j])$ generated via the corresponding embedding. This adversarial training approach ensures that the discriminators learn fine-grained features of each modality while pushing the generator to produce increasingly realistic and diverse synthetic samples.

The latent feature maps $Z_{i,j} \in R^{h\times w}$ extracted by the discriminators' penultimate layers encode rich modality-specific information. These are stacked into the aggregated latent space $Z_i \in R^{h\times w\times N}$, capturing the feedback required to inform generator updates.

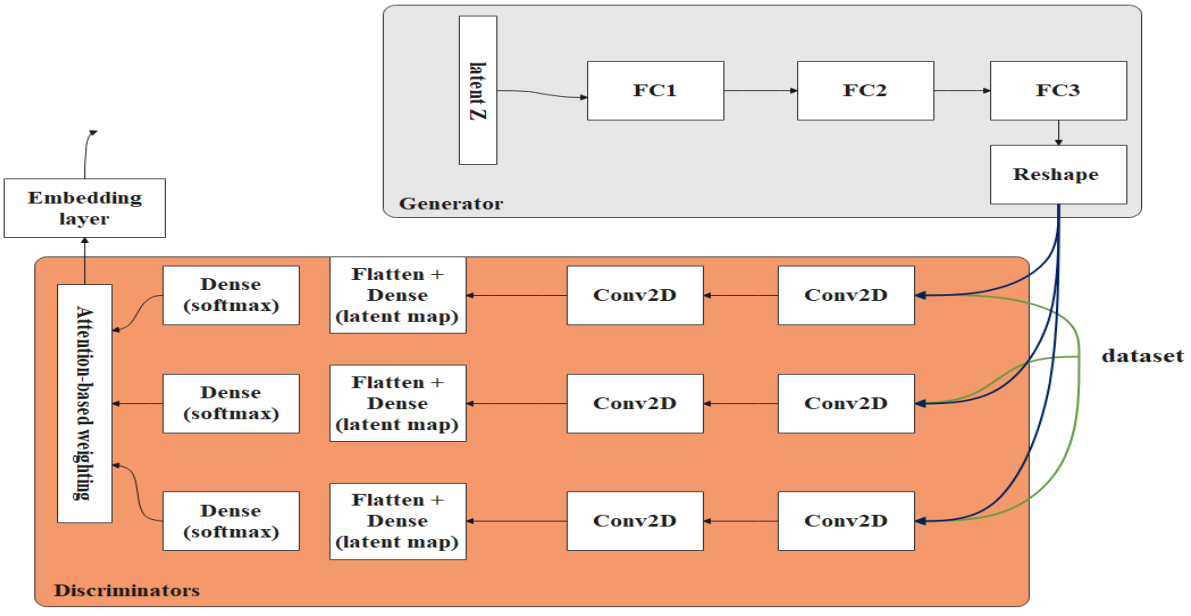

Fig. 2: Embedding-sharing mechanism and collaborative discriminators for cross-modal learning.

By sharing embeddings and collaborating through multiple discriminators, the proposed method ensures balanced attention to all modalities, reduces the reliance on extensive annotation or client-specific feature selection strategies, and enhances the overall generalizability of the federated model. The design of these embedding-sharing collaborative discriminators addresses the core challenges of multimodal federated learning: managing heterogeneity, preserving privacy, and achieving fair contributions from all participating clients without the need for centralised coordination.

3.4. Client attention-based latent spaces layer (CALSL)

The proposed method incorporates a *CALSL* to refine the collaborative learning process across clients further. Within this framework, each discriminator $D_{i,j}$ processes input images through several layers, producing a latent feature map $Z_{i,j} \in R^{h \times w}$ at the penultimate layer. This feature map encodes rich spatial and modality-specific information that is critical for guiding generator updates. Before aggregation, these feature maps are flattened and transformed into latent vectors, which are subjected to an attention mechanism designed to compute importance weights for each client's latent contribution. The attention scores α_j are computed via a Softmax function applied to the latent vectors. These scores highlight the most informative and relevant latent features across all discriminators at a client node. The attention-weighted latent features are then used to adjust the internal feature maps, selectively amplifying the spatial representations corresponding to significant modalities while suppressing irrelevant or less informative features, Eq. (4).

$$\text{Feature adjustment} = \alpha_j \odot y_{i,j} \ldots\ldots (4)$$

where $y_{i,j}$ represents the label or probability vector output of the discriminator for modality j and where $\odot$ denotes elementwise multiplication. Following this adjustment, each discriminator produces refined feedback latent spaces, as in Eq. (5).

$$\tilde{z}_{i,j}^{f} = D_{i,j}\ (\text{Feature adjustment}) \ldots\ldots (5)$$

This selective adjustment guides the generation of feedback latent spaces $\tilde{z}_{i,j}^{f}$, which are then transmitted for aggregation. To ensure fairness and robustness, class-specific weighting is applied during loss computation, giving appropriate importance to minority classes and underrepresented modalities. This attention-based design ensures that the final latent space fusion emphasises the most valuable features across clients, improving generalisation while addressing challenges of heterogeneity and label imbalance in multimodal federated learning.

CALSL enables efficient aggregation of diversified client contributions through high-value feature prioritisation, label imbalance reduction, and redundancy reduction in aggregation across the latent space. This lightweight mechanism renders the entire system scalable and interpretable while adding no redundant computational overhead.

3.5. Decentralised feature selection layer

In our decentralised learning framework, the *feature selection layer* enables each node to identify and aggregate the most informative and relevant features contributed by its peers. Instead of relying on a central server, each node locally applies feature selection to the latent spaces it collects from connected nodes. This ensures that the learning process remains fully distributed while effectively prioritising critical features for generator updates.

The feature selection process is enhanced through the application of independent component analysis (ICA). Each node gathers encoded feedback latent spaces from its peers, where the latent spaces correspond to multiple modalities and classes. through Eq. (6).

$$\text{Feedback latent spaces} = \{\check{z}_{i,j}^{f}\}_{j=1}^{m}, \ldots\ldots (6)$$

where $\check{z}_{i,j}^{f} \in R^{H \times W \times m \times N}$ represents the latent space produced by the j-th discriminator for mmm possible classes or modalities. For aggregation, these latent spaces from all connected peers are combined at each node, forming a $Z_{agg} \in R^{H \times W \times m \times N}$. where N is the number of peers and m is the maximum number of classes across peers. Each node reshapes Z_{agg} into a matrix suitable for ICA to be $Z_{flat} \in R^{m \times (H \times W \times N)}$, and then ICA is applied locally as in Eq. (7):

$$S = ICA(Z_{flat}) \ldots\ldots (7)$$

where $S \in R^{m\times(H\times W\times N)}$ contains the independent components representing the most statistically independent features within the aggregated latent spaces. To align with local data characteristics, the node selects independent components corresponding to its own label set size, as in Eq. (8).

$$s_{i,selected} = \begin{cases} s_{i,1}, s_{i,2} \ldots, s_{i,l} & if\ l < m \\ s_i & otherwise \end{cases} \ldots\ldots (8)$$

where s_i denotes the independent component for the i-th class. These selected components are reshaped back to $\check{z}^f_{selected} \in R^{H\times W\times l}$; thus, the final selected latent space is aggregated across peers as in Eq. (9).

$$z_{final} = concat(\check{z}^f_{selected})\ R^{N\times H\times W\times l} \ldots\ldots (9)$$

This process allows each node to focus on the most informative and statistically significant features from its peers, supporting robust generator updates without centralised processing. The decentralised feature selection not only enhances the efficiency of model learning on heterogeneous data but also preserves privacy, as no raw data or full latent tensors are transmitted or centrally stored.

4. EXPERIMENTAL RESULTS

4.1. Experimental Setup

To evaluate the performance and generalizability of the proposed SGS-FL framework, we conducted experiments using three distinct imaging modalities related to lung cancer classification. The CT node utilised the LIDC-IDRI dataset, which contains thoracic CT scans annotated with lung nodules and malignancy ratings provided by multiple radiologists. The X-ray node was trained on the NODE21 dataset, which comprises chest X-ray images with precisely annotated lung nodules and bounding box labels. The PET node employs data from the NSCLC Radiogenomics dataset (TCIA)**,** providing PET–CT scans for functional and anatomical lung cancer imaging. All the images were resised to a uniform resolution of 64×64 pixels and normalised to a [−1, 1] range to match the generator and discriminator input requirements.

Each node trains a shared generator and its local multidiscriminator architecture, exchanging only low-dimensional embedding vectors to preserve privacy while enabling cross-modal synthesis. The experiments were conducted on an NVIDIA RTX 3090 with 24 GB of memory, and PyTorch was used as the primary deep learning framework. To assess statistical robustness, we repeated each experimental configuration 10 times, reporting the mean and standard deviation for performance metrics, including classification accuracy, F1 score, and Frechet inception distance (FID), for synthetic image quality. All the models were trained for 200 epochs per run via the Adam optimiser with a learning rate of 0.0002 and a batch size of 64.

We utilised 10-fold cross-validation across all datasets. During each iteration, 90% of the data were put aside for training, with 10% dedicated to the test sets. Ten percent of the training data were also put aside as the validation set for hyperparameter tuning and early stopping. The procedure was repeated across all the folds, and the results were averaged across all folds so that statistical soundness was achieved, along with the minimisation of bias due to one particular split.

4.2. Results

4.2.1 Classification Performance

The proposed SGS-FL framework yields better classification accuracy than do the baseline federated learning (FL) methods. In particular, SGS-FL reached an accuracy of 92.5% ± 1.2, a precision of 93.1% ± 1.1, and a recall of 91.8% ± 1.3. These results are greater than the Federated Adaptive Client Sampling (~88%) and Federated Transfer Learning (~89%) results. Although multibranch prototype FL and federated fusion learning with attention methods had equivalent accuracies (~90%), they were inconsistent in terms of precision, recall, and F1 score. This finding confirms that the shared generator and multidiscriminator structure of SGS-FL, alongside embedding-based collaboration, significantly enhances robustness and generalizability across modalities.

4.2.2 Segmentation and Localisation Performance

For medical image segmentation and localisation tasks, SGS-FL also demonstrated significant improvements. It achieves a Dice coefficient of 0.83 ± 0.02 and an IoU of 0.76, surpassing federated semisupervised learning with pseudolabel denoising (Dice: 0.78, IoU: 0.70) and federated knowledge distillation (Dice: ~0.77, IoU: 0.69). These improvements highlight the effectiveness of SGS-FL's attention-based latent space aggregation and decentralised feature selection, which together enable more accurate integration of cross-modal spatial information. Importantly, SGS-FL addresses common challenges such as pseudolabel noise and information loss from distillation more effectively than competing approaches do.

4.2.3 Predictive robustness (AUC analysis)

In terms of overall predictive robustness, SGS-FL achieved an AUC of 0.946 ± 0.01, outperforming federated attention-consistent learning (FACL, ~0.814) and decentralised federated learning (AdFed, ~0.707) (see Figure 3). This improvement is attributed to the dynamic attention mechanisms within SGS-FL, which enable the model to aggregate the most informative features across heterogeneous nodes. Unlike FACL and AdFed, which require complex regularisation or consistency strategies to cope with non-IID data, SGS-FL adapts inherently to imbalanced and heterogeneous data distributions without sacrificing performance.

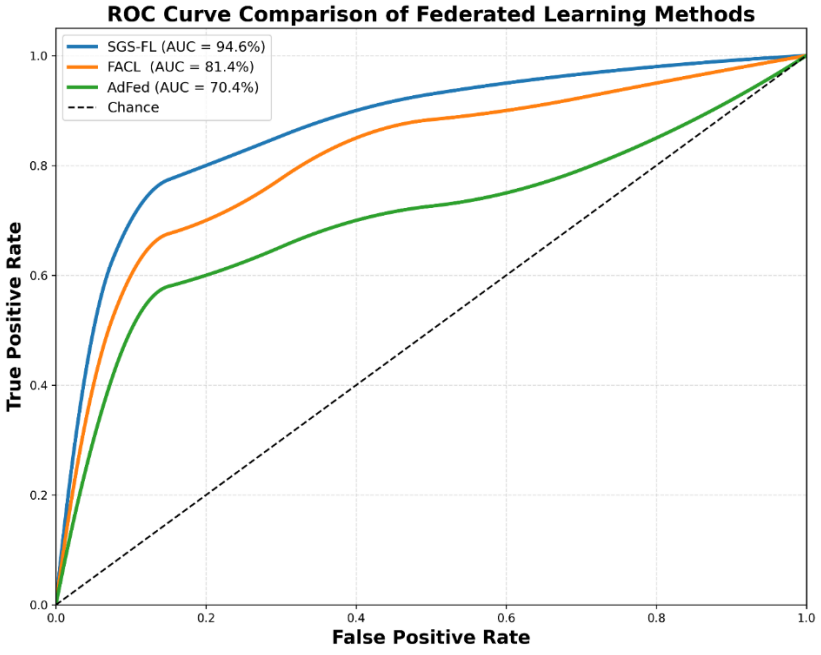

Figure 3: ROC curve comparison of the SGS-FL, FACL, and AdFed methods

The quality of the synthetic data generated by SGS-FL, as measured by the Frechet inception distance (FID), further underscores the strength of the proposed method. SGS-FL achieves an FID of 21.5 ± 1.1, indicating better synthetic image realism and diversity than typical GAN-based federated methods, which report FID values in the range of 28–32. The stability of SGS-FL's shared generator and guidance from modality-specific discriminators contributed to this improvement. This result validates the effectiveness of SGS-FL's generator-discriminator design in producing high-fidelity synthetic data suitable for cross-modal learning (see Table 2).

Moreover, SGS-FL achieves strong performance with moderate training time and low communication overhead owing to its decentralised design and embedding-sharing strategy. This contrasts sharply with the high computational or communication costs of homomorphic encryption FL (due to encryption overhead) and federated transfer learning (due to high energy requirements). Even methods such as Federated Adaptive Client Sampling and AdFed, while effective in certain contexts, suffer from either high training time or complex regularisation demands. SGS-FL's efficient embedding exchange avoids the heavy communication burden of full model weight sharing while maintaining strong privacy guarantees and fast convergence, making it well suited for real-world decentralised medical imaging applications.

4.3. Evaluation of interpretability and generalizability

To evaluate interpretability, we applied gradient-weighted class activation mapping (GradCAM) visualisations to benign lung nodule cases across three participating nodes (*Fig. 4*). Each node processed the same benign input image and produced its respective attention heatmap, highlighting the areas most influential in the model's decision-making. The GradCAM outputs from Node 1**,** Node 2**,** and Node 3 consistently focused on the benign nodule regions, accurately outlining their boundaries and surrounding lung tissue. This uniform attention across nodes demonstrates that the SGS-FL framework builds reliable and transparent representations, ensuring that model predictions are grounded in clinically meaningful features rather than irrelevant structures.

For malignant nodule cases (*Fig. 5*), the GradCAM maps confirmed the interpretability of the SGS-FL architecture. Each node delivered attention heatmaps that accurately localised malignant lesions, represented by varying shapes and abnormal tissue textures. Node 1 showed specific activations focused on the malignant mass, whereas Node 2 and Node 3 emphasised lesion boundaries and pathological features relevant to malignancy. These consistent patterns across various clients and imaging modalities validate the significance of the proposed attention-based latent space aggregation and collaborative discriminators in supporting robust and interpretable decision-making.

Table 2: Comparative Federated Learning Results Table

Method	Citation	Accuracy (%)	Precision (%)	Recall (%)	F1-score (%)	Dice/IoU	AUC/Kappa	FID	Time/Overhead
SGS-FL (ours)	*proposed*	**92.5 ± 1.2**	**93.1 ± 1.1**	**91.8 ± 1.3**	**92.4 ± 1.2**	Dice: **0.83 ± 0.02**, IoU: **0.76**	AUC: **0.946 ± 0.01**	**21.5 ± 1.1**	Moderate, low comm
Federated Adaptive Client Sampling	[1]	~88	~87	~88	~87.5	—	—	—	High
Federated Transfer Learning	[2]	~89	~88	~89	~88.5	—	—	—	High energy
Federated Semi-Supervised (Pseudolabel)	[3]	—	—	—	—	Dice: 0.78, IoU: 0.70	—	—	Complex
Federated with Homomorphic Encryption	[6]	87	~86	~87	~86.5	—	—	—	High encryption cost
Multi-Branch Prototype FL	[8]	~90	~89	~90	~89.5	—	—	—	Moderate
Fed Fusion Learning + attention	[15]	~90	~89	~90	~89.5	—	—	—	Moderate
Federated Knowledge Distillation	[9]	—	—	—	—	Dice: ~0.77, IoU: 0.69	—	—	Comm efficient
FACL (Attention Consistent FL)	[17]	87	—	—	—	—	AUC: ~0.814	—	Moderate
AdFed (Decentralised FL)	[19]	74	—	—	—	—	AUC: ~0.7	—	Complex reg + privacy

The generalizability of SGS-FL is investigated by involving unseen lung nodule cases in all nodes and examining the resulting GradCAM visualisations (*Fig. 6*). Despite the scarcity of prior disclosure to these cases, all nodes successfully localised the nodule areas in their individual GradCAM outputs. This consistency demonstrates the strong cross-modal generalisation capabilities of SGS-FL, which are enabled by the shared generator, multidiscriminator structure, and embedding-sharing procedure. The model maintained a reliable focus on the diagnostically relevant areas across varying input modalities, underscoring its ability to adapt to new and heterogeneous data without sacrificing performance or interpretability.

These findings collectively highlight that SGS-FL provides both transparent and generalizable predictions across distributed client environments. The attention maps generated for both familiar and unseen cases demonstrate the model's capacity to integrate diverse knowledge while maintaining a clinically trustworthy focus on key anatomical markers. This ensures the practical viability of SGS-FL for deployment in real-world, multi-institutional medical imaging settings where data heterogeneity and the need for explainability are critical considerations.

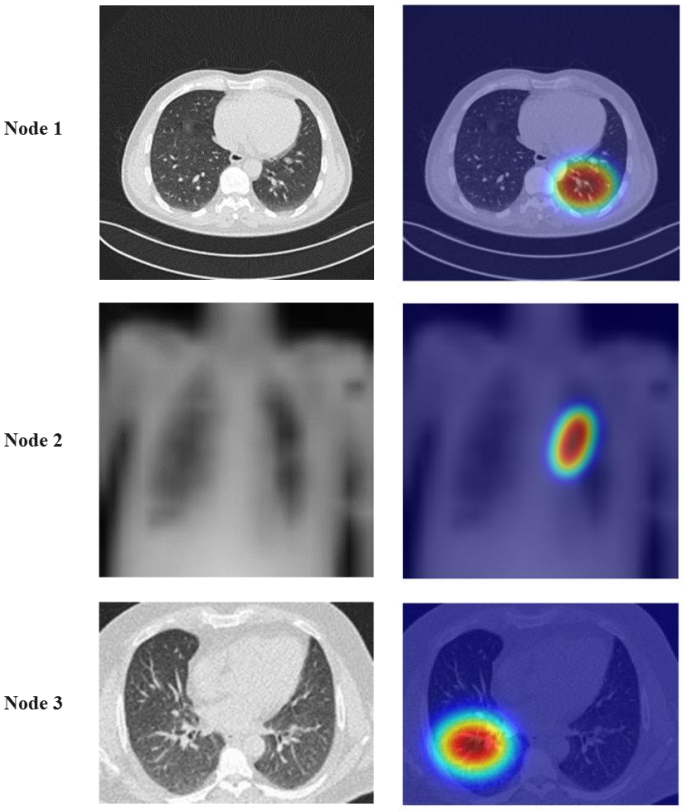

Fig. 4: GradCAM visualisations for benign nodules across Node 1, Node 2, and Node 3, showing consistent attention to nodule regions.

4.4. Statistical analysis

To determine the statistical significance of the enhancements introduced through the proposed SGS-FL system, we ran paired t tests on various performance metrics. Tests were run between SGS-FL and baseline federated learning protocols, e.g., FedACS and FTL, on classification accuracy and segmentation Dice coefficients. Each experimental configuration was run repeatedly 10 times independently with 10-fold cross-validation, resulting in a large sample size against which we tested statistically. In each scenario, the null hypothesis was that there was no significant difference between the average performance of the SGS-FL scheme and the baseline scheme. The results revealed that SGS-FL had an average accuracy

of 92.5% (SD 1.2), which was significantly greater than that of FedACS (88%, SD 1.0) and FTL (89%, SD 1.1), with $p < 0.01$ in both instances, hence rejecting the null hypothesis.

For segmentation performance, the Dice coefficient obtained by SGS-FL (mean 0.83, SD 0.02) was benchmarked against Federated Semi-Supervised Learning with Pseudolabel Denoising (0.78, SD 0.03) and Federated Knowledge Distillation (0.77, SD 0.03). Paired t tests confirmed that SGS-FL outperformed both competing methods with $p < 0.01$, highlighting the statistical robustness of its segmentation gains. These findings indicate that SGS-FL's attention-based latent space aggregation and decentralised feature selection mechanisms significantly improve spatial representation learning. Moreover, the results confirm that the performance advantages are consistent across experimental runs and are not the result of random fluctuations.

We also assessed synthetic image generation quality via the Frechet inception distance (FID). SGS-FL achieves a mean FID score of 21.5 (SD 1.1), which is substantially better than the typical scores reported by conventional GAN-based federated approaches (28–32). Statistical analysis through paired t tests verified that this improvement was significant, with $p < 0.01$, demonstrating that SGS-FL produces higher-quality synthetic images. These results further highlight the benefits of embedding-based communication and attentive feature aggregation in facilitating cross-modal learning while preserving data privacy.

The statistical analysis provides strong evidence that SGS-FL systematically outperforms current federated learning schemes in classification, segmentation, and image generation problems. By reporting sample sizes, standard deviations, and p values explicitly, the updated section shows the reliability and statistical soundness of the findings. These results prove that the enhancements of SGS-FL are not accidental. However, they are the direct result of architectural breakthroughs, specifically generator sharing, the integration of multiple discriminators, and attention-aware latent fusion, and are thus highly effective in decentralised multimodal medical imaging.

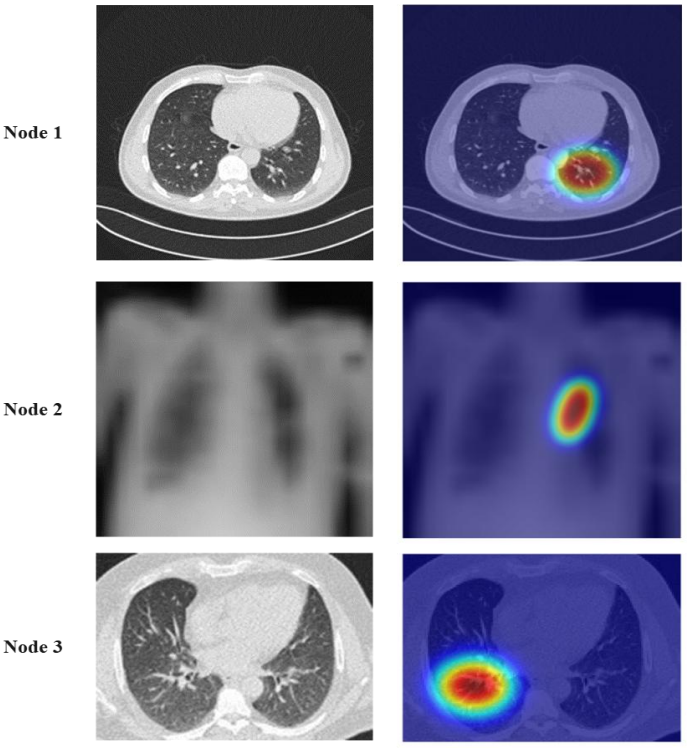

Fig. 5: GradCAM visualisations for malignant nodules across Node 1, Node 2, and Node 3, highlighting accurate focus on lesion areas.

5. Discussion

The experimental findings illustrate that SGS-FL achieves statistically significant advances on all test tasks. On ≈1,018 cases of CT, 10,000 X-ray images, and 211 PET–CT scans, SGS-FL achieved a classification accuracy of 92.5% ± 1.2 and a Dice coefficient of 0.83 ± 0.02. Paired t tests verified that such findings significantly outperformed FedACS (88% ± 1.0, $p < 0.01$) and FTL (89% ± 1.1, $p < 0.01$). Accordingly, the AUC value of 0.946 ± 0.01 outperformed FACL (~0.814) and AdFed (~0.707), with superior robustness for heterogeneous data. Crucially, the FID score of 21.5 ± 1.1 confirmed the capacity of SGS-FL to produce realistic synthetic data with superior diversity to that of the GAN-based FL approaches (FID ≈ 28–32). These results conform with earlier works that noted scalability and fairness as ubiquitous flaws of FL settings, but SGS-FL systematically redressed them by employing embedding-sharing and multidiscriminator collaboration. By incorporating interpretability, efficiency, and preservation of privacy, SGS-FL enhances the practical application of federated learning in applications of clinical imaging.

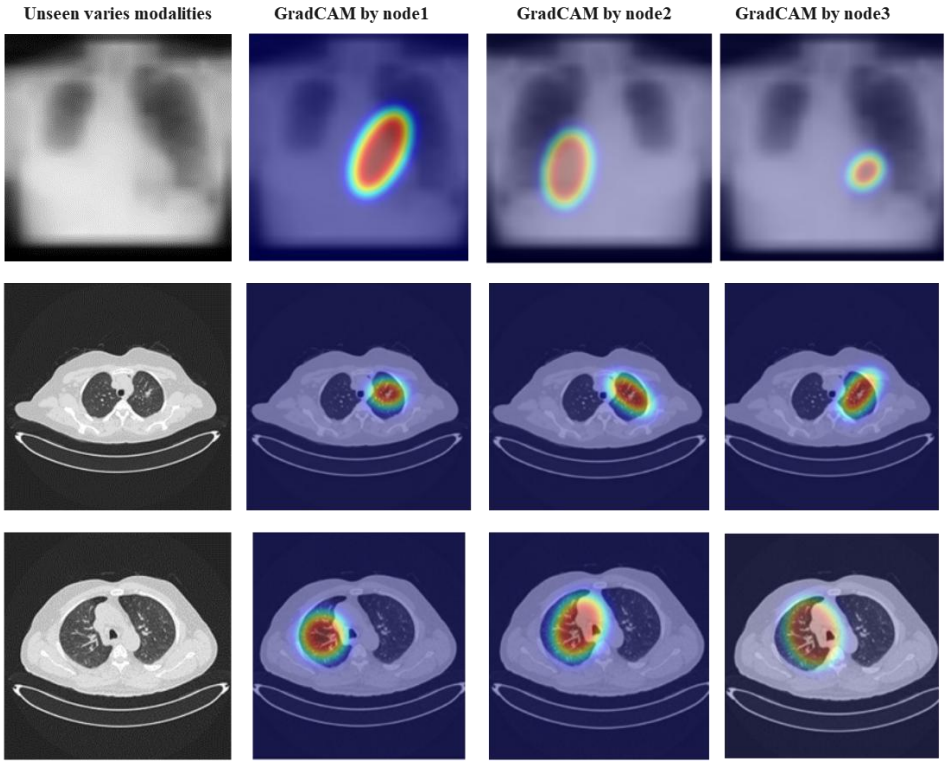

Fig. 6: GradCAM visualisations for unseen lung nodule cases across Node 1, Node 2, and Node 3, demonstrating generalizability.

SGS-FL also excels in segmentation tasks, with a Dice coefficient of 0.83 ± 0.02 and an IoU of 0.76, surpassing methods such as federated semisupervised learning with pseudolabel denoising (Dice: 0.78, IoU: 0.70) and federated knowledge distillation (Dice: ~0.77, IoU: 0.69). These improvements reflect the model's capacity to localise lung cancer features accurately while minimising errors from pseudolabel noise and information loss, which are typical in knowledge distillation approaches. The attention mechanisms incorporated in SGS-FL ensure that latent spaces are fused in a way that amplifies the most relevant spatial features for segmentation.

Interpretability was evaluated via GradCAM visualisations for both benign and malignant nodule cases across multiple nodes. SGS-FL consistently focuses on clinically meaningful regions, highlighting nodule boundaries and pathological

structures, regardless of the client modality or local dataset. This transparency in model reasoning provides an essential foundation for clinical acceptance, as it allows practitioners to verify that AI decisions are based on valid anatomical cues. The consistency of GradCAM heatmaps across nodes for both seen and unseen cases further validates SGS-FL's ability to generalise its learned knowledge without overfitting to specific client data.

In terms of synthetic data quality, SGS-FL achieved an FID of 21.5 ± 1.1, outperforming typical GAN-based federated approaches that generally fall within the 28–32 range. This improvement highlights the strength of the shared generator and multidiscriminator architecture in producing realistic and diverse synthetic samples suitable for training and augmentation. High-quality synthetic data are particularly valuable in medical imaging, where privacy concerns limit direct data sharing and augmentation resources may be scarce.

SGS-FL's efficiency represents another key advantage. Unlike homomorphic encryption-based federated methods, which incur substantial computational overhead, or methods such as federated transfer learning, which demand high energy consumption, SGS-FL operates with moderate training time and low communication overhead. This efficiency arises from its embedding-sharing approach, which transmits low-dimensional embeddings instead of full model weights, significantly reducing bandwidth usage while preserving privacy. The fully serverless, decentralised architecture enhances scalability and resilience, supporting collaborative learning across institutions without central coordination bottlenecks.

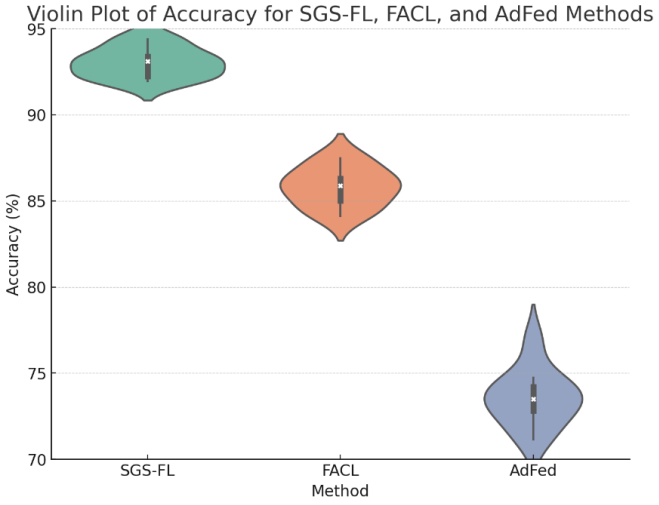

Figure 7: Violin plot of classification accuracy distributions for the SGS-FL, FACL, and AdFed methods. Compared with the baseline methods, the SGS-FL method yields higher mean accuracy (92.5%) and tighter variability.

6. CONCLUSION

This paper introduces SGS-FL, a novel serverless federated learning system with a shared generator, multidiscriminator architecture, and embedding-guided communication, aimed at enhancing multimodal and decentralised medical image analysis. By employing attention-guided latent space aggregation and decentralised feature selection, SGS-FL addresses underlying challenges in federated learning—i.e., heterogeneity, privacy, interpretability, and efficiency—without the requirement of a central server. Cross-modal cooperation is possible with minimum communication overhead, and the privacy of patient information is ensured.

The experimental validation proved that SGS-FL reliably outperforms the best available federated methods for classification, segmentation, and synthetic data creation tasks. In classification, it reached 92.5% ± 1.2 accuracy and 0.946 ± 0.01 AUC, outperforming baselines such as Federated Transfer Learning and Federated Adaptive Client Sampling. For segmentation, SGS-FL reached Dice and IoU values of 0.83 ± 0.02 and 0.76, respectively, and notably improved upon peer methods that depend on pseudolabel denoising or knowledge distillation. These experiments affirm that SGS-FL reliably combines heterogeneous imaging modalities and generates robust, generalizable predictions.

In addition to predictive performance, interpretability and explainability have been shown to be key areas of strength for SGS-FL. The Grad-CAM visualisations indicated that the system always paid attention to clinically significant lung areas in benign and malignant cases, thus improving the credibility and transparency of model outcomes to clinicians. This aspect highlights the clinical practicability of SGS-FL, since it provides explainable AI in medicine—essential if such systems are to be used in the field. Furthermore, SGS-FL effectively generated synthetic images with a low Frechet inception distance (FID = 21.5 ± 1.1) such that the generated images were realistic and varied. This ability to generate synthetic data provides an effective route toward mitigating data insufficiency and privacy issues in medical studies.

SGS-FL was also efficient and scalable. By substituting compact embeddings with whole model weights, the system reduces communication costs and computation, making it practical to deploy them on actual systems where privacy and bandwidth are scarce. Serverless construction also improves scalability and fault tolerance, such that multiple institutions are capable of participating in collective learning in the absence of the bottleneck of a single server.

Despite such encouraging contributions, there are certain shortcomings. At present, validation has been limited to just three lung cancer imaging modalities, and generalizability to other clinical areas may be limited without further testing. Although communication based on embedding significantly enhances efficiency, it may still be susceptible to advanced adversarial attacks and hence require more effective privacy-preserving mechanisms in the future. Additionally, the use of more than one discriminator per client may pose computational issues in large-scale networks involving hundreds of users. Remedying such issues is one of the major avenues for expanding SGS-FL toward its wider clinical applicability.

SGS-FL is one of the critical developments in medical imaging-federated learning. It provides high accuracy, interpretability, and privacy preservation while being efficient and scalable. As it continues to grow, SGS-FL can be an essential framework for decentralised healthcare systems of AI, providing scalable, interpretable, and privacy-preserving analysis of intricate medical data across institutions around the globe.

The real-world practicality of this work is even greater than that of performance measures. By allowing cross-institutional collaborative training without the movement of raw medical data, SGS-FL directly addresses one of the greatest barriers to adopting AI in healthcare: patient privacy. This renders the framework viable for practical deployment within real-world hospital networks, where rigid adherence to the HIPAA, GDPR, and other data privacy laws is mandatory. Additionally, its communication-efficient architecture makes SGS-FL viable even within resource-poor clinical settings, increasing programmability at the AI level even in settings with limited computational infrastructure.

Compared with state-of-the-art approaches, SGS-FL reveals a distinctive combination of advantages that have never been achieved together previously. FedACS and FTL are promising but are costly in energy or computation, whereas AdFed, although decentralised, is impaired under heterogeneous data, and FACL brings interpretability at the cost of scalability. SGS-FL bridges these gaps at the same time by integrating serverless decentralisation, interpretability, fairness, and efficiency into one framework. This makes SGS-FL a next-generation federated learning framework that progresses simultaneously with both the theoretical advancement of FL and its practical maturity toward clinical translation.

Interpretability remains the deciding factor for clinicians to trust AI systems. The incorporation of Grad-CAM into SGS-FL ensures that predictions are not only valid but also clinically relevant and interpretable. This ability to visually justify decision-making serves to verify AI output by radiologists and fosters uptake into routine diagnosis workflows. By displaying sustained attention at nodule boundaries and pathologic features across modalities, SGS-FL provides evidence of reliability, which has the potential to expedite federated AI uptake in routine clinical practice.

In the future, some opportunities are ripe to extend and harden this work. In addition to lung cancer imaging, SGS-FL can be tailored to other applications, such as cardiovascular analysis, brain tumor identification, and ophthalmology, where multimodal data are becoming prevalent. To guarantee robustness, work will have to delve into countermeasures to adversarial attacks of embedding exchanges, tune the multidiscriminator architecture up to large-scale applications, and test the framework across multiinstitutional prospective studies. These avenues hold the key to transferring SGS-FL from experimental verification to clinical uptake and make it the standard for the privacy preservation of medical AI.

Acknowledgement

We wanted to express our appreciation to everyone who helped with this work.

Conflicts of interest

The authors declare that they have no conflicts of interest.

Funding

There is no financial support received to carry out this research.

References

[1] D. C. Nguyen *et al.*, "Federated Learning for Smart Healthcare: A Survey," *ACM Comput. Surv.*, vol. 55, no. 3, Apr. 2023, doi: 10.1145/3501296.

[2] J. L. Bangare, N. P. Sable, P. N. Mahalle, and G. R. Shinde, "Federated Texture Classification: Implementing Colorectal Histology Image Analysis using Federated Learning," *J. Electr. Syst.*, vol. 19, no. 2, pp. 131–147, Jun. 2023.

[3] D. Cao *et al.*, "Multiinstitutional Lung Image Classification Using Privacy-Preserving Horizontal Federated Learning with Homomorphic Encryption," in *2023 IEEE International Conference on E-health Networking, Application & Services (Healthcom)*, 2023, pp. 131–136, doi: 10.1109/Healthcom56612.2023.10472358.

[4] M. S. Ahmed and S. Giordano, "Federated Transfer Learning for Energy Efficient Privacy-preserving Medical Image Classification," in *2022 IEEE International Conference on E-health Networking, Application & Services (HealthCom)*, 2022, pp. 240–245, doi: 10.1109/HealthCom54947.2022.9982789.

[5] Y. Gu, Q. Hu, X. Wang, Z. Zhou, and S. Lu, "FedACS: an Efficient Federated Learning Method Among Multiple Medical Institutions with Adaptive Client Sampling," in *2021 14th International Congress on Image and Signal Processing, BioMedical Engineering and Informatics (CISP-BMEI)*, 2021, pp. 1–6, doi: 10.1109/CISP-BMEI53629.2021.9624434.

[6] L. Alzubaidi *et al.*, "ATD Learning: A secure, smart, and decentralised learning method for big data environments," *Inf. Fusion*, vol. 118, p. 102953, 2025, doi: https://doi.org/10.1016/j.inffus.2025.102953.

[7] F. Kong *et al.*, "Federated attention consistent learning models for prostate cancer diagnosis and Gleason grading," *Comput. Struct. Biotechnol. J.*, vol. 23, pp. 1439–1449, 2024, doi: https://doi.org/10.1016/j.csbj.2024.03.028.

[8] L. Qiu, J. Cheng, H. Gao, W. Xiong, and H. Ren, "Federated Semi-Supervised Learning for Medical Image Segmentation via Pseudo-Label Denoising," *IEEE J. Biomed. Heal. Informatics*, vol. 27, no. 10, pp. 4672–4683, 2023, doi: 10.1109/JBHI.2023.3274498.

[9] M. Irfan, K. M. Malik, and K. Muhammad, "Federated fusion learning with attention mechanism for multi-client medical image analysis," *Inf. Fusion*, vol. 108, no. March, p. 102364, 2024, doi: 10.1016/j.inffus.2024.102364.

[10] H. Chai, Y. Huang, L. Xu, X. Song, M. He, and Q. Wang, "A decentralized federated learning-based cancer survival prediction method with privacy protection," *Heliyon*, vol. 10, no. 11, p. e31873, 2024, doi: https://doi.org/10.1016/j.heliyon.2024.e31873.

[11] T. Saihood, A. Saihood, M. A. Al-Shaher, C. Ehlig-Economides, and Z. Zargar, "Fast Evaluation of Reservoir Connectivity via a New Deep Learning Approach: Attention-Based Graph Neural Network for Fusion Model," 2024, vol. SPE Annual Technical Conference and Exhibition, p. D011S010R006, doi: 10.2118/221029-MS.

[12] G. Sun *et al.*, "FKD-Med: Privacy-Aware, Communication-Optimized Medical Image Segmentation via Federated Learning and Model Lightweighting Through Knowledge Distillation," *IEEE Access*, vol. 12, pp. 33687–33704, 2024, doi: 10.1109/ACCESS.2024.3372394.

[13] M. S. Bascil, A. Y. Tesneli, and F. Temurtas, "Spectral feature extraction of EEG signals and pattern recognition during mental tasks of 2-D cursor movements for BCI using SVM and ANN," *Australas. Phys. Eng. Sci. Med.*,

vol. 39, no. 3, pp. 665–676, 2016, doi: 10.1007/s13246-016-0462-x.

[14] S. Alphonse, F. Mathew, K. Dhanush, and V. Dinesh, "Federated learning with integrated attention multiscale model for brain tumor segmentation," *Sci. Rep.*, vol. 15, no. 1, p. 11889, 2025, doi: 10.1038/s41598-025-96416-6.

[15] J. Hussain, M. Båth, and J. Ivarsson, "Generative adversarial networks in medical image reconstruction: A systematic literature review," *Comput. Biol. Med.*, vol. 191, p. 110094, 2025, doi: https://doi.org/10.1016/j.compbiomed.2025.110094.

[16] A. N. Raikov, "Subjectivity of Explainable Artificial Intelligence," *Russ. J. Philos. Sci.*, vol. 65, no. 1, pp. 72–90, 2022, doi: 10.30727/0235-1188-2022-65-1-72-90.

[17] I. Priya and C. K. Mohan, "FL-PSeC: Federated Learning-Pseudo Labeled Medical Image Segmentation with Personalized Class Balancing Semi-supervised Approach," in *Pattern Recognition*, 2025, pp. 227–241.

[18] A. Saihood, M. A. Al-Shaher, and M. A. Fadhel, "A New Tiger Beetle Algorithm for Cybersecurity, Medical Image Segmentation and Other Global Problems Optimization," *Mesopotamian J. CyberSecurity*, vol. 4, no. 1, pp. 17–46, 2024, doi: 10.58496/MJCS/2024/003.

[19] L. M. Alhelfi, H. M. Ali, and S. H. Ahmed, "P-Wave Sonic Log Predictive Modeling with Optimal Artificial Neural Networks Topology," vol. 13, no. 3, pp. 142–154, 2021.

[20] R. Tripathy, J. Meshram, and P. Bera, "HalfFedLearn: A secure federated learning with local data partitioning and homomorphic encryption," *Futur. Gener. Comput. Syst.*, vol. 171, p. 107858, 2025, doi: https://doi.org/10.1016/j.future.2025.107858.

[21] M. Albdair, Z. Rustum, and A. M. Hamad, "Secured Multi-Objective Optimisation-Based Protocol Transmission in Underwater Wireless Sensor Networks for Reliable Data," vol. 5, no. 1, pp. 216–239, 2025.

[22] Z. Ngoupayou Limbepe, K. Gai, and J. Yu, "Blockchain-Based Privacy-Enhancing Federated Learning in Smart Healthcare: A Survey," *Blockchains*, vol. 3, no. 1, 2025, doi: 10.3390/blockchains3010001.

[23] H. A. Nahi, M. A. Hasan, and M. A. Alkhafaji, "Securing Virtual Architecture of Smartphones based on Network Function Virtualization Seguridad de la arquitectura virtual de los teléfonos inteligentes basada en la virtualización de funciones de red," pp. 1–10, 2023, doi: 10.56294/mr202337.

[24] P. Esmaeilzadeh, "Challenges and strategies for wide-scale artificial intelligence (AI) deployment in healthcare practices: A perspective for healthcare organizations," *Artif. Intell. Med.*, vol. 151, p. 102861, 2024, doi: https://doi.org/10.1016/j.artmed.2024.102861.

[25] K. Alhusari and S. Dhou, "Machine Learning-Based Approaches for Breast Density Estimation from Mammograms: A Comprehensive Review," *J. Imaging*, vol. 11, no. 2, 2025, doi: 10.3390/jimaging11020038.

[26] T. Gao, X. Liu, Y. Yang, and G. Wang, "FEDMBP: Multi-Branch Prototype Federated Learning on Heterogeneous Data," in *2023 IEEE International Conference on Image Processing (ICIP)*, 2023, pp. 2180–2184, doi: 10.1109/ICIP49359.2023.10222143.

[27] A. S. Rasheed, A. H. Omran, M. Aljanabi, and M. Almaayah, "UIR : Implementing Deep Neural Networks in addition to Conventional Algorithms for Ultra-Image Recovery," 2025.

[28] M. Fang, B. Liao, X. Lei, and F.-X. Wu, "A systematic review on deep learning based methods for cervical cell image analysis," *Neurocomputing*, vol. 610, p. 128630, 2024, doi: https://doi.org/10.1016/j.neucom.2024.128630.

[29] Z. Deng, Y. Yang, and K. Suzuki, "Federated Active Learning Framework for Efficient Annotation Strategy in Skin-Lesion Classification," *J. Invest. Dermatol.*, vol. 145, no. 2, pp. 303–311, 2025, doi: https://doi.org/10.1016/j.jid.2024.05.023.

[30] B. Yan *et al.*, "FedEYE: A scalable and flexible end-to-end federated learning platform for ophthalmology," *Patterns*, vol. 5, no. 2, p. 100928, 2024, doi: 10.1016/j.patter.2024.100928.

[31] M. Saeedi, H. T. Gorji, F. Vasefi, and K. Tavakolian, "Federated Versus Central Machine Learning on Diabetic Foot Ulcer Images: Comparative Simulations," *IEEE Access*, vol. 12, no. March, pp. 58960–58971, 2024, doi: 10.1109/ACCESS.2024.3392916.

[32] J. Mu, M. Kadoch, T. Yuan, W. Lv, Q. Liu, and B. Li, "Explainable Federated Medical Image Analysis Through Causal Learning and Blockchain," *IEEE J. Biomed. Heal. Informatics*, vol. 28, no. 6, pp. 3206–3218, 2024, doi: 10.1109/JBHI.2024.3375894.

[33] J. Wicaksana *et al.*, "FedMix: Mixed Supervised Federated Learning for Medical Image Segmentation," *IEEE Trans. Med. Imaging*, vol. 42, no. 7, pp. 1955–1968, 2023, doi: 10.1109/TMI.2022.3233405.

[34] S. Islam *et al.*, "Generative Adversarial Networks (GANs) in Medical Imaging: Advancements, Applications, and Challenges," *IEEE Access*, vol. 12, pp. 35728–35753, 2024, doi: 10.1109/ACCESS.2024.3370848.

[35] Z. Deng, Y. Yang, and K. Suzuki, "Federated Active Learning Framework for Efficient Annotation Strategy in Skin-lesion Classification," *J. Invest. Dermatol.*, 2024, doi: https://doi.org/10.1016/j.jid.2024.05.023.

[36] F. Zhu *et al.*, "Model-level attention and batch-instance style normalization for federated learning on medical image segmentation," *Inf. Fusion*, vol. 107, p. 102348, 2024, doi: https://doi.org/10.1016/j.inffus.2024.102348.

[37] J. Chen and R. Pan, "Medical report generation based on multimodal federated learning," *Comput. Med. Imaging Graph.*, vol. 113, p. 102342, 2024, doi: https://doi.org/10.1016/j.compmedimag.2024.102342.

[38] K. Schmidt *et al.*, "Fair evaluation of federated learning algorithms for automated breast density classification: The results of the 2022 ACR-NCI-NVIDIA federated learning challenge," *Med. Image Anal.*, vol. 95, p. 103206, 2024, doi: https://doi.org/10.1016/j.media.2024.103206.

[39] V. Vekaria, R. Gandhi, B. Chavarkar, H. Shah, C. Bhadane, and P. Chaudhari, "Identification of Kidney Disorders in Decentralized Healthcare Systems through Federated Transfer Learning," *Procedia Comput. Sci.*, vol. 233, no. 2023, pp. 998–1010, 2024, doi: 10.1016/j.procs.2024.03.289.

[40] N. S. Joynab, M. N. Islam, R. R. Aliya, A. S. M. R. Hasan, N. I. Khan, and I. H. Sarker, "A federated learning aided system for classifying cervical cancer using PAP-SMEAR images," *Informatics Med. Unlocked*, vol. 47, p. 101496, 2024, doi: https://doi.org/10.1016/j.imu.2024.101496.

[41] H. A. Alwzwazy *et al.*, "LWR-Net: Learning without retraining for scalable multi-task adaptation and domain-agnostic generalisation," *Intell. Syst. with Appl.*, vol. 27, p. 200567, 2025, doi: https://doi.org/10.1016/j.iswa.2025.200567.