FEDQUADCONTRA: AMBIGUITY-AWARE CONTRAPOSITIVE QUADRUPLET LEARNING FOR FEDERATED REMOTE SENSING SCENE CLASSIFICATION

Anonymous authors
Paper under double-blind review

ABSTRACT

Federated learning enables collaborative model training across decentralized data while preserving data locality, yet its performance can degrade substantially under statistical heterogeneity, where skewed local class distributions induce biased representations and inconsistent local optimization. This challenge is particularly pronounced in remote sensing scene classification, where high inter-class similarity and substantial intra-class variability further complicate discriminative representation learning. We introduce FedQuadContra, an ambiguity-aware contrapositive quadruplet learning framework for federated representation learning under heterogeneous client distributions. Unlike conventional pairwise and triplet metric-learning objectives that capture a limited set of inter-sample relationships, FedQuadContra constructs each quadruplet using an anchor, a same-class positive, a hard negative, and an ambiguity-aware contrapositive negative selected from a semantic class distinct from both the anchor and hard-negative classes. This construction extends metric learning beyond a single competing negative class by explicitly modeling multiple inter-class relationships within each quadruplet. The resulting multi-relational objective jointly promotes intra-class compactness, hard inter-class separation, and discrimination among representation-wise confusable samples, enabling more discriminative representation learning under heterogeneous client distributions. We evaluate FedQuadContra under varying degrees of statistical heterogeneity on remote sensing scene classification benchmarks and across multiple network architectures. Extensive experiments show that FedQuadContra consistently outperforms representative federated learning and federated contrastive-learning methods, including the recent RS-CCL approach, across the evaluated data distributions, with particularly pronounced improvements under severe heterogeneity. These results demonstrate the effectiveness of ambiguity-aware multi-relational representation learning for federated remote sensing scene classification.

1 INTRODUCTION

Federated learning (FL) enables multiple clients to collaboratively learn a global model while keeping their data decentralized McMahan et al. (2017). This paradigm is particularly attractive in settings where data sharing is constrained by privacy, ownership, bandwidth, or institutional boundaries. However, the effectiveness of FL is strongly affected by statistical heterogeneity across clients. In practice, local data are often non-independent and identically distributed (Non-IID), with substantial differences in class composition and sample frequencies across clients. Consequently, local optimization can be biased toward client-specific distributions, leading to inconsistent local representations and degradation of the aggregated global model. Methods such as FedProx Li et al. (2020) and SCAFFOLD Karimireddy et al. (2020) address this problem primarily by regularizing local optimization or correcting client drift, but learning discriminative representations under severe data heterogeneity remains challenging.

This challenge is especially relevant to remote sensing (RS) scene classification. Images belonging to different scene categories can exhibit high visual similarity, while images within the same category may vary considerably because of differences in spatial arrangement, scale, appearance,

Reviewers: please read the Reviewer Guidelines (iclr.cc/Conferences/2027/ReviewerGuidelines) and the AI Policy for Reviewers (iclr.cc/Conferences/2027/AIPolicyForReviewers). If you used AI to expand, edit, or polish your review, please provide the input text to the LLM. Better still, consider skipping the LLM and submitting your original text: we, and the authors, are much more interested in your unedited thoughts than in what an LLM has to say. AI-assisted or not, you are putting your name and reputation behind your review: LLM-generated falsehoods, hallucinations or misrepresentations are subject to disciplinary action, which may include desk-rejecting all papers you have authored.

and acquisition conditions. When such data are distributed across clients, Non-IID partitioning can further restrict the diversity of semantic relationships observed during local training. A client may therefore learn a representation space dominated by its locally available classes, even though the resulting model must ultimately discriminate among a substantially richer set of categories after federated aggregation. Recent studies have consequently investigated FL specifically for RS scene classification, including representation- and vision-language-based approaches Lin et al. (2025); Chen et al. (2026).

Representation learning provides a complementary direction for addressing heterogeneity in FL. Model-contrastive approaches such as MOON Li et al. (2021a) encourage consistency between local and global model representations, while prototype-based approaches exploit class-level representations to guide local learning Mu et al. (2023); Long et al. (2022). More recently, RS-CCL Chen et al. (2026) couples supervised classification and contrastive feature learning for federated RS image classification. These methods demonstrate the importance of representation-level constraints in heterogeneous FL. Nevertheless, most existing contrastive or metric-learning formulations are primarily organized around positive–negative relationships, prototype alignment, or local–global representation consistency. Such formulations do not explicitly model the richer relational structure that can arise when multiple semantically distinct classes occupy competing neighborhoods in the embedding space.

We argue that this limitation is particularly important in multi-class federated learning. Consider an anchor sample and a same-class positive. Conventional triplet learning additionally identifies a hard negative from a different class and optimizes the relative distances among these three samples. Although effective, this construction characterizes inter-class ambiguity primarily through a single competing negative class. In a heterogeneous multi-class setting, however, an anchor may be simultaneously confusable with samples from several distinct semantic classes. Restricting the learning objective to a single negative relationship may therefore leave additional ambiguous inter-class structure insufficiently constrained. This motivates a broader question: *can local representation learning exploit multiple competing class relationships to learn more discriminative representations under federated data heterogeneity?*

To address this question, we propose **FedQuadContra**, an **ambiguity-aware contrapositive quadruplet learning** framework for federated representation learning. FedQuadContra extends conventional triplet metric learning by constructing a quadruplet consisting of an anchor, a same-class positive, a hard negative, and an *ambiguity-aware contrapositive negative* selected from a semantic class distinct from both the anchor and hard-negative classes. The additional sample is therefore not merely another negative instance: together with the hard negative, it introduces an additional competing class relationship into the local embedding space. FedQuadContra exploits this structure through a multi-relational objective that simultaneously encourages anchor–positive compactness, anchor–hard-negative separation, anchor–contrapositive separation, and discrimination between the hard-negative and contrapositive relationships. This formulation exposes local learning to multiple class-diverse relational constraints within each quadruplet and is designed to improve discriminative representation learning when client distributions are heterogeneous. To address this question, we propose **FedQuadContra**, an **ambiguity-aware contrapositive quadruplet learning** framework for federated representation learning. FedQuadContra extends conventional triplet metric learning by constructing a quadruplet consisting of an anchor, a same-class positive, a hard negative, and an *ambiguity-aware contrapositive negative* selected from a semantic class distinct from both the anchor and hard-negative classes. The additional sample is therefore not merely another negative instance: together with the hard negative, it introduces an additional competing class relationship into the local embedding space. FedQuadContra exploits this structure through a multi-relational objective that simultaneously encourages anchor–positive compactness, anchor–hard-negative separation, anchor–contrapositive separation, and discrimination between the hard-negative and contrapositive relationships. This formulation exposes local learning to multiple class-diverse relational constraints within each quadruplet and is designed to improve discriminative representation learning when client distributions are heterogeneous.

An overview of the motivation and the proposed FedQuadContra framework is illustrated in Fig. 1.

FedQuadContra is integrated into the standard federated optimization process without requiring clients to exchange raw samples or intermediate representations. Each participating client performs local representation and classification learning using its private data, after which model parameters

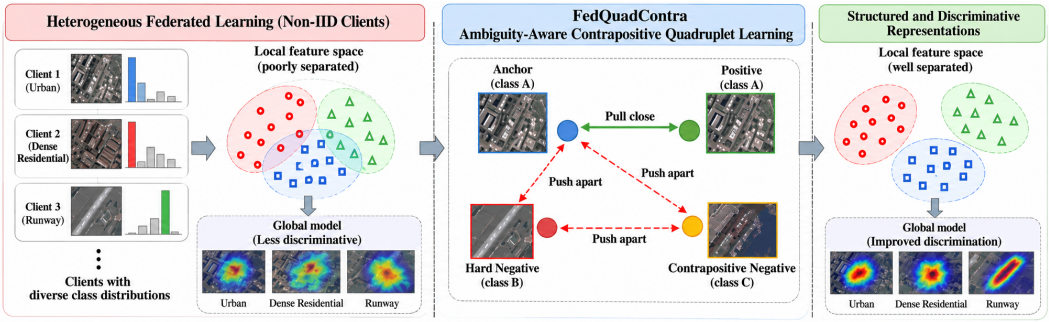

Figure 1: Overview of FedQuadContra under heterogeneous federated learning. FedQuadContra uses ambiguity-aware contrapositive quadruplets to capture multiple competing inter-class relationships, encouraging more discriminative representations under Non-IID client distributions.

are aggregated to update the global model. Importantly, the proposed quadruplet objective operates at the representation level and is not tied to a particular network architecture. We therefore evaluate FedQuadContra across multiple backbone architectures and under varying degrees of statistical heterogeneity. Experiments on established RS scene classification benchmarks with IID and Dirichlet-based Non-IID partitions show consistent improvements over representative optimization-, contrastive-, and representation-based FL methods, including the recent RS-CCL approach. The improvements are particularly evident under highly heterogeneous client distributions, supporting the use of class-diverse relational constraints for federated representation learning.

The main contributions of this work are summarized as follows:

- We introduce **FedQuadContra**, an ambiguity-aware contrapositive quadruplet learning framework for discriminative representation learning under heterogeneous federated data distributions.
- We propose a **class-diverse quadruplet construction** that augments the conventional anchor–positive–hard-negative relation with an ambiguity-aware contrapositive negative from a semantic class distinct from both the anchor and hard-negative classes. This enables the local learner to capture multiple competing inter-class relationships rather than relying on a single negative class.
- We formulate a **multi-relational quadruplet objective** that jointly promotes intra-class compactness and discrimination across anchor–hard-negative, anchor–contrapositive, and hard-negative–contrapositive relationships, providing richer local representation constraints under client heterogeneity.
- We conduct extensive experiments under IID and multiple levels of Non-IID data heterogeneity and across different network architectures. The results demonstrate consistent improvements over representative FL and federated representation-learning approaches, including RS-CCL, particularly under severe statistical heterogeneity.

2 Related Work

2.1 Federated Learning under Data Heterogeneity

Federated learning (FL) enables multiple clients to collaboratively train a global model while retaining their data locally McMahan et al. (2017). A fundamental challenge in FL is statistical heterogeneity, where differences in local class distributions can induce inconsistent client updates and degrade global generalization. Several approaches have addressed this problem from an optimization perspective. FedProx Li et al. (2020) introduces a proximal regularizer to restrict excessive deviation of local models from the global model, SCAFFOLD Karimireddy et al. (2020) employs control variates to correct client drift, and FedNova Wang et al. (2020) addresses objective inconsistency arising from heterogeneous local computation. FedDyn Acar et al. (2021) dynamically regularizes local objectives to improve consistency between local and global optimization, while

FedBN Li et al. (2021b) retains local batch-normalization statistics to address feature-distribution shifts across clients.

More recent approaches address heterogeneity at the representation level. MOON Li et al. (2021a) employs model-level contrastive learning to encourage consistency between local and global representations. FedDecorr Shi et al. (2023) mitigates dimensional collapse Jing et al. (2022)through feature decorrelation, while FedRCL Seo et al. (2024) develops a relaxed supervised contrastive objective for heterogeneous client distributions. These studies demonstrate that addressing representation quality can complement conventional optimization-based mechanisms for heterogeneous FL.

2.2 Federated Representation and Metric Learning

Federated representation learning has increasingly exploited class-level and feature-level relationships to improve learning under heterogeneous distributions. FedProc Mu et al. (2023) employs global class prototypes to guide local representation learning, whereas FedPCL Long et al. (2022) uses prototype-wise contrastive learning for knowledge sharing across clients. FedCDG Yu et al. (2024) further considers representation learning for improving generalization across heterogeneous client distributions. Together with MOON, FedDecorr, and FedRCL, these methods highlight the importance of explicitly structuring the representation space in federated learning. Supervised contrastive learning further exploits label information to encourage same-class compactness and different-class separation in the embedding space Khosla et al. (2020), providing an important basis for representation-level learning in supervised settings.

Deep metric learning structures the embedding space by enforcing relational constraints among samples, ranging from pairwise contrastive objectives to triplet and higher-order metric formulations Hadsell et al. (2006); Schroff et al. (2015); Sohn (2016); Song et al. (2016); Chen et al. (2017); Yu & Tao (2019). Pairwise contrastive learning encourages similar samples to remain close while separating dissimilar samples, whereas triplet learning jointly considers an anchor, a same-class positive, and a different-class negative. Higher-order metric objectives extend this principle by incorporating additional sample relationships to provide richer structural constraints on the embedding spaceSun et al. (2020); Wang et al. (2019).

Building on this perspective, FedQuadContra introduces *ambiguity-aware multi-relational representation learning* for heterogeneous federated distributions. It constructs a class-diverse quadruplet comprising an anchor, a same-class positive, a hard negative, and an ambiguity-aware contrapositive negative from a semantic class distinct from both the anchor and hard-negative classes. The resulting objective explicitly constrains the anchor–hard-negative, anchor–contrapositive, and hard-negative–contrapositive relationships, thereby modeling multiple competing semantic relationships during local representation learning.

2.3 Federated Learning for Remote Sensing Scene Classification

Federated learning has gained increasing attention in remote sensing (RS), where imagery is often distributed across institutions, geographic regions, or decentralized archives. Youssef et al. Youssef et al. (2024) investigate FL for remote sensing scene classification, while Buyuktas et al. study collaborative learning across decentralized RS image archives. More specialized approaches include FedRSCLIP Lin et al. (2025), which employs shared and private vision–language prompts for global–local knowledge alignment, and Fed-RSSC , which explores a semi-decentralized framework for RS scene classification. More recently, RS-CCL Chen et al. (2026) couples classification and supervised contrastive learning to improve representations under heterogeneous client distributions. Despite these advances, explicit modeling of multiple competing semantic relationships within local metric learning remains underexplored. FedQuadContra addresses this gap through ambiguity-aware, class-diverse quadruplet construction and multi-relational representation learning.

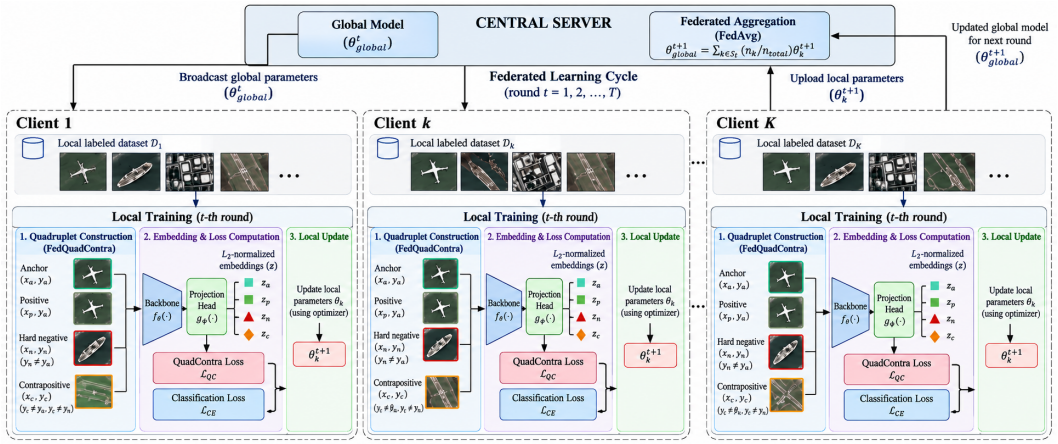

Figure 2: Overall framework of FedQuadContra. At each communication round, the global model is distributed to participating clients, where ambiguity-aware class-diverse quadruplets are constructed from the local labeled data. FedQuadContra jointly optimizes supervised classification and multi-relational quadruplet objectives over ℓ_2-normalized embeddings using squared Euclidean distance. The locally updated models are subsequently aggregated at the server to obtain the global model for the next round.

3 Proposed FedQuadContra Framework

3.1 Federated Learning Problem Formulation

We consider a supervised federated learning (FL) system consisting of a central server and K clients. Each client k maintains a private labeled dataset $\mathcal{D}_k = \{(x_i, y_i)\}_{i=1}^{N_k}$, where N_k denotes the number of local training samples. The raw client data remain local throughout training, while model parameters are exchanged with the central server.

The objective is to learn a global model with parameters $\mathbf{w}$ by minimizing the weighted combination of local objectives:

$$\min_{\mathbf{w}} F(\mathbf{w}) = \sum_{k=1}^{K} \frac{N_k}{N} F_k(\mathbf{w}), \qquad N = \sum_{k=1}^{K} N_k, \tag{1}$$

where $F_k(\mathbf{w})$ denotes the supervised learning objective on $\mathcal{D}_k$.

We focus on statistically heterogeneous client distributions, where $P_k(X, Y) \neq P_j(X, Y)$ for different clients k and j. In particular, class-distribution skew causes clients to observe different class proportions and, potentially, different subsets of semantic classes. Such heterogeneity can lead to client-specific representation spaces and inconsistent local updates before aggregation. FedQuadContra addresses this setting by introducing ambiguity-aware, multi-relational representation learning into the local supervised objective.

3.2 FedQuadContra Framework

Figure 2 illustrates the overall FedQuadContra framework. At communication round t, the central server broadcasts the current global model $\mathbf{w}^t$ to the selected clients. Each participating client initializes its local model with the received parameters and performs supervised training using its private labeled dataset $\mathcal{D}_k$. FedQuadContra modifies the local representation-learning process while retaining the standard federated parameter-exchange mechanism.

For an input image x_i, the feature encoder $f_\theta(\cdot)$ extracts a representation that is mapped by the projection head $g_\phi(\cdot)$ into a d-dimensional embedding spaceCollins et al. (2021); Chen et al. (2020). The projected representation is ℓ_2-normalized as

$$z_i = \frac{g_\phi(f_\theta(x_i))}{\|g_\phi(f_\theta(x_i))\|_2}, \qquad D(z_i, z_j) = \|z_i - z_j\|_2^2, \tag{2}$$

where $D(\cdot, \cdot)$ denotes the squared Euclidean distance between normalized embeddings. Since $\|z_i\|_2 = 1$, this distance is independent of embedding magnitude and provides a consistent measure for ambiguity-aware sample selection and relational learning.

Using the local labeled data, FedQuadContra constructs class-diverse quadruplets consisting of an anchor, a same-class positive, a hard negative, and an ambiguity-aware contrapositive negative from a semantic class distinct from both the anchor and hard-negative classes. The resulting embeddings are optimized using the proposed multi-relational QuadContra loss, while the classification branch is trained using the standard supervised cross-entropy loss. Thus, local training jointly learns class-discriminative predictions and explicitly structures relationships among semantically competing samples.

After completing the prescribed local training epochs, each participating client returns its updated model parameters $\mathbf{w}_k^{t+1}$ to the server. The server performs data-size-weighted aggregation to obtain the global model for the next communication round. FedQuadContra therefore introduces ambiguity-aware relational learning entirely within local training and does not require the exchange of raw images, embeddings, prototypes, or class-specific statistics among clients.

3.3 Ambiguity-Aware Contrapositive Quadruplet Construction

Conventional triplet learning represents inter-class ambiguity using a single negative class. Under heterogeneous federated distributions, however, an anchor may be close to samples from multiple semantic classes within a local embedding space. To capture such class-diverse ambiguity, FedQuadContra constructs a quadruplet consisting of an anchor x_a, a same-class positive x_p, a hard negative x_n, and an ambiguity-aware contrapositive negative x_c:

$$\mathcal{Q}_i = (x_a, x_p, x_n, x_c), \qquad y_a = y_p, \quad y_n \neq y_a, \quad y_c \neq y_a, \quad y_c \neq y_n. \tag{3}$$

Thus, each valid quadruplet contains samples from at least three semantic classes. The anchor–positive pair provides the intra-class reference, whereas x_n and x_c represent two distinct competing classes.

Hard-negative selection. For each anchor, a positive sample is selected from the same class. Among the locally available samples belonging to classes different from the anchor, the hard negative is selected as the closest sample in the normalized embedding space:

$$n^* = \arg \min_{j:\, y_j \neq y_a} D(z_a, z_j). \tag{4}$$

The selected x_n therefore represents the strongest immediate different-class competitor to the anchor according to the squared Euclidean distance defined in Eq. 2He et al. (2020).

Ambiguity-aware contrapositive selection. A single hard negative captures ambiguity with respect to only one competing class. FedQuadContra therefore identifies an additional confusable sample while explicitly excluding both the anchor class and the already selected hard-negative class:

$$c^* = \arg \min_{j:\, y_j \neq y_a,\, y_j \neq y_n} D(z_a, z_j). \tag{5}$$

The resulting x_c is the closest eligible sample from a semantic class distinct from both y_a and y_n. Hence, the contrapositive sample is not an arbitrary second negative; it represents an additional representation-wise ambiguous class that is not captured by the hard negative alone.

This construction is particularly relevant under Non-IID class distributions, where different clients may observe different class compositions and consequently different ambiguous class neighborhoods. Whenever at least two negative classes are locally available for an anchor, FedQuadContra exploits two distinct competing classes through x_n and x_c. If an eligible third-class sample is unavailable, a contrapositive quadruplet is not constructed for that anchor, thereby preserving the class-diversity constraint. Overall, the construction captures the anchor–positive, anchor–hard-negative, anchor–contrapositive, and hard-negative–contrapositive relationships within a single class-diverse quadruplet.

3.4 MULTI-RELATIONAL QUADCONTRA LEARNING

Given a valid quadruplet $\mathcal{Q} = (x_a, x_p, x_n, x_c)$, FedQuadContra jointly optimizes the relationships among the corresponding normalized embeddings (z_a, z_p, z_n, z_c). The anchor–positive distance serves as the intra-class reference, while the hard-negative and contrapositive samples represent two distinct competing semantic classes. The proposed QuadContra loss is defined as

$$\begin{aligned}\mathcal{L}_{\mathrm{QC}} = \frac{1}{|\mathbb{Q}|} \sum_{\mathcal{Q}\in\mathbb{Q}} \Big[& [D(z_a, z_p) - D(z_a, z_n) + \alpha]_+ \\ & + [D(z_a, z_p) - D(z_n, z_c) + \beta]_+ \\ & + [D(z_a, z_p) - D(z_a, z_c) + \gamma]_+ \Big],\end{aligned} \tag{6}$$

where $[u]_+ = \max(0, u)$, $\mathbb{Q}$ denotes the set of valid local quadruplets, and α, β, and γ are positive margin parameters.

The three terms impose complementary relational constraints. The first separates the anchor from its hard negative relative to the same-class anchor–positive distance, thereby promoting intra-class compactness and hard inter-class discrimination. The second explicitly separates the hard negative from the contrapositive sample; since $y_n \neq y_c$, this term models the relationship between two distinct competing semantic classes rather than treating x_c merely as another anchor-centered negative. The third separates the anchor from the ambiguity-aware contrapositive sample, thereby addressing an additional confusable class relationship. Together, these constraints jointly organize the anchor–positive, anchor–hard-negative, anchor–contrapositive, and hard-negative–contrapositive relationships within the local embedding space.

The QuadContra loss is jointly optimized with the standard supervised cross-entropy loss. To additionally limit excessive local model deviation from the current global model under heterogeneous client distributions, we include proximal regularization. The complete local objective for client k at communication round t is

$$\mathcal{L}_k = \mathcal{L}_{\mathrm{CE}}^k + \lambda_{\mathrm{QC}} \mathcal{L}_{\mathrm{QC}}^k + \frac{\mu}{2} \left\| \mathbf{w}_k - \mathbf{w}^t \right\|_2^2, \tag{7}$$

where λ_{QC} controls the contribution of the proposed QuadContra loss and μ determines the strength of proximal regularization. Thus, the classification loss provides class-level supervision, the QuadContra loss structures multiple competing semantic relationships in the local embedding space, and the proximal term constrains local parameter drift from the global model.

3.5 THEORETICAL ANALYSIS

The multi-relational structure of FedQuadContra provides stronger separation constraints than an anchor-centered metric objective, as formalized below.

Multi-Relational Separation: For a FedQuadContra quadruplet (z_a, z_p, z_n, z_c), if $\mathcal{L}_{\mathrm{QC}} = 0$, then

$$\begin{aligned} D(z_a, z_n) &\geq D(z_a, z_p) + \alpha, \\ D(z_n, z_c) &\geq D(z_a, z_p) + \beta, \\ D(z_a, z_c) &\geq D(z_a, z_p) + \gamma. \end{aligned} \tag{8}$$

Thus, relative to the intra-class anchor–positive distance, FedQuadContra simultaneously enforces separation between the anchor and two distinct competing classes, as well as between the competing classes themselves.

Proof. Each term in Eq. 6 is a non-negative hinge loss. Therefore, $\mathcal{L}_{\mathrm{QC}} = 0$ implies that all three hinge terms are individually zero. Rearranging the corresponding margin constraints directly yields Eq. 8. □

Proposition 3.5 highlights the role of the contrapositive sample. While the first inequality provides the conventional anchor–hard-negative separation, the second and third introduce additional constraints involving a distinct competing class. In particular, the $D(z_n, z_c)$ constraint explicitly structures the relationship between the two competing classes, distinguishing the proposed objective from simply using an additional anchor-centered negative.

Table 1: Classification accuracy (%) on UC Merced Land Use under varying levels of Dirichlet Non-IID data skew (α) and the IID setting. Results are reported as mean $\pm$ standard deviation. The best results are shown in bold.

Backbone	Method	$\alpha = 0.05$	$\alpha = 0.1$	$\alpha = 0.5$	IID
ResNet-18	FedAvg McMahan et al. (2017)	79.75 ± 2.2	80.16 ± 2.0	81.50 ± 1.8	82.74 ± 1.0
	FedProx Li et al. (2020)	80.22 ± 1.5	80.53 ± 2.3	81.22 ± 1.4	82.20 ± 0.8
	SCAFFOLD Karimireddy et al. (2020)	79.97 ± 2.4	80.99 ± 1.9	85.23 ± 1.1	85.09 ± 1.1
	MOON Li et al. (2021a)	80.75 ± 1.7	81.66 ± 2.0	86.07 ± 1.7	86.11 ± 0.5
	FedPCL Long et al. (2022)	80.06 ± 2.0	81.25 ± 1.5	85.21 ± 1.5	85.35 ± 1.0
	FedCDG Yu et al. (2024)	80.53 ± 2.3	81.65 ± 2.0	85.97 ± 1.0	85.86 ± 0.7
	FedProc Mu et al. (2023)	78.22 ± 2.5	79.55 ± 2.2	82.34 ± 1.1	82.76 ± 1.2
	FedDecorr Shi et al. (2023)	82.08 ± 1.5	82.24 ± 1.7	86.15 ± 1.5	86.22 ± 0.6
	FedDyn Acar et al. (2021)	82.19 ± 2.1	83.33 ± 1.9	86.27 ± 1.7	86.13 ± 0.9
	FedRCL Seo et al. (2024)	83.50 ± 2.3	84.29 ± 1.9	86.97 ± 0.6	88.00 ± 0.9
	RS-CCL Chen et al. (2026)	87.17 ± 1.5	88.07 ± 1.5	90.56 ± 0.8	90.82 ± 0.7
	FedQuadContra (Ours)	$\mathbf{93.10 \pm 1.1}$	$\mathbf{95.71 \pm 0.8}$	$\mathbf{96.81 \pm 0.5}$	$\mathbf{97.62 \pm 0.2}$

Table 2: Classification accuracy (%) on NWPU-RESISC45 under varying levels of Dirichlet Non-IID data skew (α) and the IID setting. Results are reported as mean $\pm$ standard deviation.

Backbone	Method	$\alpha = 0.05$	$\alpha = 0.1$	$\alpha = 0.5$	IID
ResNet-18	FedAvg McMahan et al. (2017)	78.34 ± 1.7	80.95 ± 0.5	82.78 ± 0.4	83.53 ± 0.4
	FedProx Li et al. (2020)	79.07 ± 2.6	81.55 ± 1.1	82.82 ± 0.3	83.12 ± 0.4
	SCAFFOLD Karimireddy et al. (2020)	80.01 ± 2.9	81.09 ± 2.1	83.87 ± 0.9	84.24 ± 0.6
	MOON Li et al. (2021a)	81.95 ± 1.8	82.88 ± 2.0	84.04 ± 1.5	84.35 ± 0.8
	FedPCL Long et al. (2022)	80.54 ± 1.3	81.79 ± 1.5	83.89 ± 1.5	84.14 ± 0.7
	FedCDG Yu et al. (2024)	82.00 ± 2.7	82.84 ± 2.5	84.17 ± 1.2	84.50 ± 1.2
	FedProc Mu et al. (2023)	77.97 ± 2.0	80.99 ± 0.9	83.82 ± 0.8	84.22 ± 0.8
	FedDecorr Shi et al. (2023)	80.00 ± 1.7	81.93 ± 1.5	84.23 ± 0.6	84.35 ± 0.7
	FedDyn Acar et al. (2021)	80.26 ± 2.5	82.08 ± 1.9	84.60 ± 1.6	84.32 ± 0.7
	FedRCL Seo et al. (2024)	82.44 ± 2.9	83.26 ± 2.5	85.83 ± 1.3	86.05 ± 0.5
	RS-CCL Chen et al. (2026)	85.04 ± 2.0	86.67 ± 1.1	89.48 ± 1.0	89.96 ± 0.2
	FedQuadContra (Ours)	$\mathbf{88.31 \pm 1.9}$	$\mathbf{91.21 \pm 1.3}$	$\mathbf{92.51 \pm 0.6}$	$\mathbf{93.25 \pm 0.3}$

4 Experimental Results

Datasets and Federated Settings. We evaluate FedQuadContra on two standard remote sensing scene classification benchmarks: UCM Yang & Newsam (2010), containing 2,100 images from 21 classes, and NWPU-RESISC45 Cheng et al. (2017), containing 31,500 images from 45 classes. UCM is partitioned across 42 clients with 10 clients participating per communication round (24%), whereas NWPU-RESISC45 is distributed across 100 clients with 10 active clients per round (10%). Statistical heterogeneity is simulated using Dirichlet class partitioning with $\alpha \in \{0.05, 0.1, 0.5\}$, where smaller α represents stronger class-distribution skew. An IID partition is additionally considered as a reference setting.

4.1 Implementation Details

We use ResNet-18 initialized with ImageNet weights as the backbone for all experiments. The feature encoder is followed by an MLP projection head that maps the extracted features to a 128-dimensional embedding space, followed by ℓ_2 normalization. A linear classifier maps the projected representations to the corresponding scene classes. Models are optimized using SGD with momentum 0.9 and weight decay 10^{-4}. The initial learning rate is 0.03, with exponential decay factors of 0.98 and 0.985 for NWPU-RESISC45 and UCM, respectively. The proximal regularization coefficient is fixed at $\mu = 0.01$. For NWPU-RESISC45, local models are trained for 3 epochs per communication round with a batch size of 32 for 200 rounds. For UCM, we use 5 local epochs, a batch size of 10, and 250 communication rounds. Unless otherwise stated, the same federated partitions and training configurations are used across the compared methods.

4.2 Comparison with State-of-the-Art Methods

We compare FedQuadContra with representative federated optimization and representation-learning methods, including FedAvg, FedProx, SCAFFOLD, MOON, FedPCL, FedCDG, FedProc, FedDecorr, FedDyn, FedRCL, and RS-CCL. Tables 1 and 2 report classification accuracy under three Non-IID settings and IID partitioning. FedQuadContra consistently achieves the highest accuracy on both datasets. Compared with RS-CCL, it improves NWPU-RESISC45 by 2.96–4.54 percentage points and UCM by 5.93–7.64 points across the Non-IID settings, while also achieving gains of 3.29 and 6.80 points under IID, respectively. These results demonstrate the effectiveness of multi-relational representation learning across varying levels of client heterogeneity.

4.3 Effect of Data Heterogeneity

As expected, classification performance generally decreases as the Dirichlet concentration parameter becomes smaller, reflecting the increased difficulty caused by highly skewed local class distributions. Nevertheless, FedQuadContra maintains strong performance across all evaluated levels of heterogeneity. On NWPU-RESISC45, accuracy increases from 88% at $\alpha = 0.05$ to 93.25% under IID partitioning, while UCM improves from 93.10% to 97.62%. More importantly, FedQuadContra retains a consistent advantage over the compared methods throughout this range, providing evidence that the proposed multi-relational constraints remain effective under both severe and moderate client heterogeneity.

4.4 Ablation Study

Table 3: Ablation Study on UC Merced Land Use (ResNet-18, $\alpha = 0.05$)

Configuration	**Test Accuracy (%)**
CE + FedAvg (Baseline)	79.75 ± 2.2
Triplet + FedAvg	82.00 ± 1.8
Triplet + FedProx	84.81 ± 1.4
QuadContra + FedAvg	89.20 ± 1.2
QuadContra + FedProx (Ours)	$\mathbf{93.10 \pm 1.1}$

To assess the contribution of each component, we conduct an ablation study on UCM with ResNet-18 under severe heterogeneity ($\alpha = 0.05$). Starting from FedAvg with cross-entropy (79.75%), adding Triplet loss improves accuracy to 82.00%. QuadContra with FedAvg substantially increases performance to 89.20%, outperforming Triplet+FedProx (84.81%). The complete FedQuadContra framework, combining QuadContra with proximal regularization, achieves the best performance of $93.10 \pm 1.1\%$. These results demonstrate the contribution of the proposed QuadContra objective and its complementary effect with proximal regularization.

5 Conclusion

This work introduced **FedQuadContra**, an ambiguity-aware multi-relational representation learning framework for federated remote sensing scene classification under heterogeneous client distributions. FedQuadContra constructs class-diverse quadruplets consisting of an anchor, a same-class positive, a hard negative, and an ambiguity-aware contrapositive negative from a distinct competing class. The proposed three-term QuadContra objective explicitly models the anchor–hard-negative, anchor–contrapositive, and hard-negative–contrapositive relationships, providing richer local representation constraints while retaining standard federated parameter aggregation. Experiments on UCM and NWPU-RESISC45 under different levels of Dirichlet data heterogeneity demonstrate consistent improvements over representative federated optimization and representation-learning approaches, including RS-CCL. The results support the effectiveness of ambiguity-aware, multi-relational representation learning for improving discriminative scene representations in heterogeneous federated settings. Future work will investigate adaptive contrapositive mining and extend the framework to larger-scale and multimodal remote sensing data.

AI USE STATEMENT

In this work, we used generative AI tools for language editing, improving the clarity and organization of the manuscript, and assisting with LaTeX formatting. We have not used generative AI tools for generating experimental results, performing experiments, or drawing scientific conclusions. Additionally, we used generative AI tools for assisting with literature organization and manuscript presentation. We have reviewed all AI-assisted work. We take responsibility for the final content of this work, including text, claims, and artifacts produced with the aid of generative AI.

REFERENCES

Durmus Alp Emre Acar, Yue Zhao, Ramon Matas Navarro, Matthew Mattina, Paul N. Whatmough, and Venkatesh Saligrama. Federated learning based on dynamic regularization. In *International Conference on Learning Representations (ICLR)*, 2021.

Ting Chen, Simon Kornblith, Mohammad Norouzi, and Geoffrey Hinton. A simple framework for contrastive learning of visual representations. In *International Conference on Machine Learning (ICML)*, pp. 1597–1607. PMLR, 2020.

Weihua Chen, Xiaotang Chen, Jianguo Zhang, and Kaiqi Huang. Beyond triplet loss: A deep quadruplet network for person re-identification. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 403–412, 2017.

Xiaoyu Chen, Weiying Xie, Xin Zhang, Hangyu Ye, Yunsong Li, and Leyuan Fang. Federated coupled contrastive learning on remote sensing image classification. *IEEE Transactions on Geoscience and Remote Sensing*, 64, 2026. doi: 10.1109/TGRS.2026.3654941.

Gong Cheng, Junwei Han, and Xiaoqiang Lu. Remote sensing image scene classification: Benchmark and state of the art. *Proceedings of the IEEE*, 105(10):1865–1883, 2017.

Liam Collins, Hamed Hassani, Aryan Mokhtari, and Sanjay Shakkottai. Exploiting shared representations for personalized federated learning. In *International Conference on Machine Learning (ICML)*, pp. 2089–2099. PMLR, 2021.

Raia Hadsell, Sumit Chopra, and Yann LeCun. Dimensionality reduction by learning an invariant mapping. In *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 1735–1742, 2006. doi: 10.1109/CVPR.2006.100.

Kaiming He, Haoqi Fan, Yuxin Wu, Saining Xie, and Ross Girshick. Momentum contrast for unsupervised visual representation learning. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 9729–9738, 2020.

Li Jing, Pascal Vincent, Yann LeCun, and Yuandong Tian. Understanding dimensional collapse in contrastive self-supervised learning. In *International Conference on Learning Representations (ICLR)*, 2022.

Sai Praneeth Karimireddy, Satyen Kale, Mehryar Mohri, Sashank Reddi, Sebastian Stich, and Ananda Theertha Suresh. SCAFFOLD: Stochastic controlled averaging for federated learning. In *Proceedings of the 37th International Conference on Machine Learning*, volume 119, pp. 5132–5143, 2020.

Prannay Khosla, Piotr Teterwak, Chen Wang, Aaron Sarna, Yonglong Tian, Phillip Isola, Aaron Maschinot, Ce Liu, and Dilip Krishnan. Supervised contrastive learning. In *Advances in Neural Information Processing Systems*, volume 33, pp. 18661–18673, 2020.

Qinbin Li, Bingsheng He, and Dawn Song. Model-contrastive federated learning. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 10713–10722, 2021a. doi: 10.1109/CVPR46437.2021.01057.

Tian Li, Anit Kumar Sahu, Manzil Zaheer, Maziar Sanjabi, Ameet Talwalkar, and Virginia Smith. Federated optimization in heterogeneous networks. In *Proceedings of Machine Learning and Systems*, volume 2, pp. 429–450, 2020.

Xiaoxiao Li, Meirui Jiang, Xiaofei Zhang, Michael Kamp, and Qi Dou. Fedbn: Federated learning on non-iid features via local batch normalization. In *International Conference on Learning Representations*, 2021b.

Hui Lin, Chao Zhang, Danfeng Hong, Kexin Dong, and Congcong Wen. FedRSCLIP: Federated learning for remote sensing scene classification using vision-language models. *IEEE Geoscience and Remote Sensing Magazine*, 13(3):260–275, 2025. doi: 10.1109/MGRS.2025.3556532.

Guodong Long, Jie Ma, Jing Jiang, Lu Liu, Tianyi Zhou, and Yue Tan. Fedpcl: Learning to blend representations for federated prototype learning. In *Proceedings of the AAAI Conference on Artificial Intelligence*, 2022.

H. Brendan McMahan, Eider Moore, Daniel Ramage, Seth Hampson, and Blaise Agüera y Arcas. Communication-efficient learning of deep networks from decentralized data. In *Proceedings of the 20th International Conference on Artificial Intelligence and Statistics*, volume 54, pp. 1273–1282, 2017.

Xutong Mu, Yulong Shen, Ke Cheng, Xueli Geng, Jiaxuan Fu, Tao Zhang, and Zhiwei Zhang. Fedproc: Prototypical contrastive federated learning on non-iid data. *Future Generation Computer Systems*, 143:93–104, 2023. doi: 10.1016/j.future.2023.01.019.

Florian Schroff, Dmitry Kalenichenko, and James Philbin. Facenet: A unified embedding for face recognition and clustering. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 815–823, 2015. doi: 10.1109/CVPR.2015.7298682.

Seonguk Seo, Jihong Kim, Geeho Kim, and Bohyung Han. Relaxed contrastive learning for federated learning. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 12279–12288, 2024.

Yujun Shi, Jian Liang, Wenqing Zhang, Vincent Y. F. Tan, and Song Bai. Feddecorr: Tackling the dimensional collapse in federated learning. In *International Conference on Learning Representations (ICLR)*, 2023.

Kihyuk Sohn. Improved deep metric learning with multi-class n-pair loss objective. In *Advances in Neural Information Processing Systems*, volume 29, pp. 1857–1865, 2016.

Hyun Oh Song, Yu Xiang, Stefanie Jegelka, and Silvio Savarese. Deep metric learning via lifted structured feature embedding. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 4004–4012, 2016.

Yifan Sun, Changmao Cheng, Yuhan Zhang, Chi Zhang, Liang Zheng, Zhongdao Wang, and Yichen Wei. Circle loss: A unified perspective of pair similarity optimization. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 6398–6407, 2020.

Jianyu Wang, Qinghua Liu, Hao Liang, Gauri Joshi, and H. Vincent Poor. Tackling the objective inconsistency problem in heterogeneous federated optimization. In *Advances in Neural Information Processing Systems*, volume 33, pp. 7611–7623, 2020.

Xun Wang, Xintong Han, Weilin Huang, Dengke Dong, and Matthew R Scott. Multi-similarity loss with general pair weighting for deep metric learning. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 5022–5030, 2019.

Yi Yang and Shawn Newsam. Bag-of-visual-words and spatial extensions for land-use classification. In *Proceedings of the 18th SIGSPATIAL International Conference on Advances in Geographic Information Systems*, pp. 270–279, 2010.

Belgacem Ben Youssef, Lamyaa Alhmidi, Yakoub Bazi, and Mansour Zuair. Federated learning approach for remote sensing scene classification. *Remote Sensing*, 16(12):2194, 2024. doi: 10.3390/rs16122194.

Baosheng Yu and Dacheng Tao. Deep metric learning with tuplet margin loss. In *Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV)*, pp. 6490–6499, 2019.

Xinhui Yu, Dan Wang, Martin J. McKeown, and Z. Jane Wang. Contrastive-enhanced domain generalization with federated learning. *IEEE Transactions on Artificial Intelligence*, 5(4):1525–1532, 2024. doi: 10.1109/TAI.2023.3298297.

6 APPENDIX

Algorithm 1 Federated Quadruplet Contrapositive Learning (QuadContra)

Require: Total clients K, Communication rounds T, Local epochs E, Batch size B, Learning rate η, Proximal term μ, Contrastive weight λ, Margins m_{std}, m_{contra}.

Server Initialization: Initialize global model weights w_0.
for each communication round $t = 1, 2, \ldots, T$ **do**
  Sample a subset of clients $S_t \subseteq \{1, 2, \ldots, K\}$
  for each client $k \in S_t$ **in parallel do**
    $w_t^k \leftarrow \text{ClientUpdate}(k, w_{t-1})$
  end for
  Global Aggregation:
  $w_t \leftarrow \sum_{k \in S_t} \frac{n_k}{n_{S_t}} w_t^k$ ▷ n_k is the number of samples on client k
end for

function CLIENTUPDATE(k, w_{global})
  Initialize local weights $w \leftarrow w_{global}$
  for local epoch $e = 1, 2, \ldots, E$ **do**
    for batch (X, Y) in local dataset $\mathcal{D}_k$ **do**
      1. Forward Pass & L2 Normalization:
      Extract projected features: $H = \text{Projector}(\text{Backbone}(X; w))$
      Normalize to unit hypersphere: $Z = \frac{H}{\|H\|_2}$
      Compute classification logits: $\hat{Y} = \text{Classifier}(Z; w)$

      2. Objective Function Computation:
      Cross-Entropy Baseline: $\mathcal{L}_{CE} = \text{CrossEntropy}(Y, \hat{Y})$
      Mine hardest triplets contrapositives within batch (Euclidean L_2 metric):
      $\mathcal{L}_{QC} \leftarrow 0$
      for each anchor $z_a \in Z$ with label y_a **do**
        $z_p \leftarrow \arg\max_{z \in Z, y = y_a} \|z_a - z\|_2$ ▷ Hardest Positive
        $z_n \leftarrow \arg\min_{z \in Z, y \neq y_a} \|z_a - z\|_2$ ▷ Nearest Negative
        $z_c \leftarrow \arg\max_{z \in Z, y \notin \{y_a, y_n\}} \text{CosineSim}(z_a, z)$ ▷ Hardest Contrapositive
        $\mathcal{L}_{trip} = \max(0, \|z_a - z_p\|_2 - \|z_a - z_n\|_2 + m_{std})$
        $\mathcal{L}_{contra} = \max(0, \|z_a - z_p\|_2 - \|z_a - z_c\|_2 + m_{contra})$
        $\mathcal{L}_{QC} \leftarrow \mathcal{L}_{QC} + (\mathcal{L}_{trip} + \mathcal{L}_{contra})$
      end for
      Average $\mathcal{L}_{QC}$ over batch valid valid quadruplets.

      3. Proximal Regularization & Update:
      $\mathcal{L}_{prox} = \frac{\mu}{2} \|w - w_{global}\|_2^2$ ▷ FedProx Constraint
      $\mathcal{L}_{Total} = \mathcal{L}_{CE} + \lambda \mathcal{L}_{QC} + \mathcal{L}_{prox}$
      $w \leftarrow w - \eta \nabla \mathcal{L}_{Total}$
    end for
  end for
  return updated local weights w
end function

Table 4: Performance Comparison on NWPU-RESISC45 (VGG16) under varying levels of Dirichlet non-IID data skew (α) and IID settings.

Backbone	Method	$\alpha = 0.05$	$\alpha = 0.1$	$\alpha = 0.5$	IID
VGG16	FedAvg	77.85 ± 2.1	80.42 ± 1.3	82.13 ± 0.8	83.07 ± 0.6
	FedProx	78.63 ± 2.5	81.02 ± 1.4	82.25 ± 0.6	82.83 ± 0.5
	SCAFFOLD	79.18 ± 2.8	80.65 ± 2.1	83.14 ± 1.1	83.76 ± 0.7
	MOON	80.53 ± 1.8	82.03 ± 2.2	83.57 ± 1.3	84.02 ± 0.8
	FedPCL	79.93 ± 2.2	81.29 ± 2.0	83.23 ± 1.5	83.22 ± 1.1
	FedCDG	80.69 ± 2.0	81.99 ± 1.1	83.59 ± 1.3	84.20 ± 1.1
	FedProc	77.43 ± 2.3	80.18 ± 1.1	83.04 ± 0.9	83.72 ± 0.8
	FedDecorr	79.34 ± 1.9	81.52 ± 1.7	83.92 ± 0.8	84.13 ± 0.6
	FedDyn	79.62 ± 2.7	81.83 ± 2.1	84.13 ± 1.6	84.05 ± 0.9
	FedRCL	81.53 ± 2.9	82.83 ± 2.4	85.16 ± 1.2	85.64 ± 0.7
	RS-CCL (TGRS 2026)	84.03 ± 2.0	85.57 ± 1.2	88.63 ± 1.0	89.14 ± 0.6
	QuadContra (Ours)	$\mathbf{86.06 \pm 1.6}$	$\mathbf{88.08 \pm 0.9}$	$\mathbf{92.19 \pm 0.6}$	$\mathbf{93.30 \pm 0.3}$

Table 5: Performance Comparison on UC Merced Land Use (VGG16) under varying levels of Dirichlet non-IID data skew (α) and IID settings.

Backbone	Method	$\alpha = 0.05$	$\alpha = 0.1$	$\alpha = 0.5$	IID
VGG16	FedAvg	78.90 ± 2.3	79.40 ± 2.1	80.80 ± 1.9	82.10 ± 1.1
	FedProx	79.40 ± 1.6	79.80 ± 2.4	80.50 ± 1.5	81.60 ± 0.9
	SCAFFOLD	78.80 ± 2.5	80.10 ± 2.0	84.50 ± 1.2	84.30 ± 1.2
	MOON	79.60 ± 1.8	80.70 ± 2.1	85.30 ± 1.8	85.40 ± 0.6
	FedPCL	79.24 ± 1.9	80.11 ± 1.5	85.17 ± 1.2	85.20 ± 1.1
	FedCDG	79.97 ± 1.6	80.54 ± 1.8	85.36 ± 1.1	85.25 ± 0.9
	FedProc	77.50 ± 2.6	78.80 ± 2.3	81.60 ± 1.2	82.00 ± 1.3
	FedDecorr	81.30 ± 1.6	81.50 ± 1.6	85.40 ± 1.1	85.50 ± 0.7
	FedDyn	81.40 ± 2.2	82.50 ± 2.0	85.60 ± 1.8	85.40 ± 1.0
	FedRCL	82.70 ± 2.4	83.50 ± 2.0	86.20 ± 1.7	87.20 ± 1.3
	RS-CCL (TGRS 2026)	86.00 ± 1.8	86.90 ± 1.6	89.30 ± 0.9	89.50 ± 1.0
	QuadContra (Ours)	$\mathbf{88.23 \pm 1.4}$	$\mathbf{92.86 \pm 1.1}$	$\mathbf{94.85 \pm 0.7}$	$\mathbf{96.31 \pm 0.5}$

6.1 Ambiguity-Aware Contrapositive Sample Selection

For each anchor x_a, we first select a hard negative x_n as the closest sample belonging to a class different from the anchor:

$$n^* = \arg \min_{j:\, y_j \neq y_a} D(z_a, z_j). \tag{9}$$

To capture an additional competing semantic relationship, the contrapositive sample x_c is selected as the closest sample to the anchor from a class different from both the anchor and hard-negative classes:

$$c^* = \arg \min_{j:\, y_j \neq y_a,\ y_j \neq y_n} D(z_a, z_j), \tag{10}$$

where $D(z_i, z_j) = \|z_i - z_j\|_2^2$ denotes the squared Euclidean distance between ℓ_2-normalized embeddings. Thus,

$$y_c \neq y_a, \qquad y_c \neq y_n, \tag{11}$$

ensuring that (x_a, x_p, x_n, x_c) contains at least three semantic classes. Unlike an arbitrary second negative, x_c represents the most confusable eligible sample from an additional competing class. If no such third-class sample is available locally, a valid quadruplet is not formed for that anchor.

6.2 Architecture-Agnostic Evaluation: VGG16 Performance

To demonstrate that the stabilizing properties of our proposed framework are architecture-agnostic, we extend our evaluation to the VGG16 backbone on both the UC Merced and NWPU-RESISC45 benchmarks. The comparative results across varying degrees of Dirichlet non-IID data skew are reported in Table 5 and Table 4.

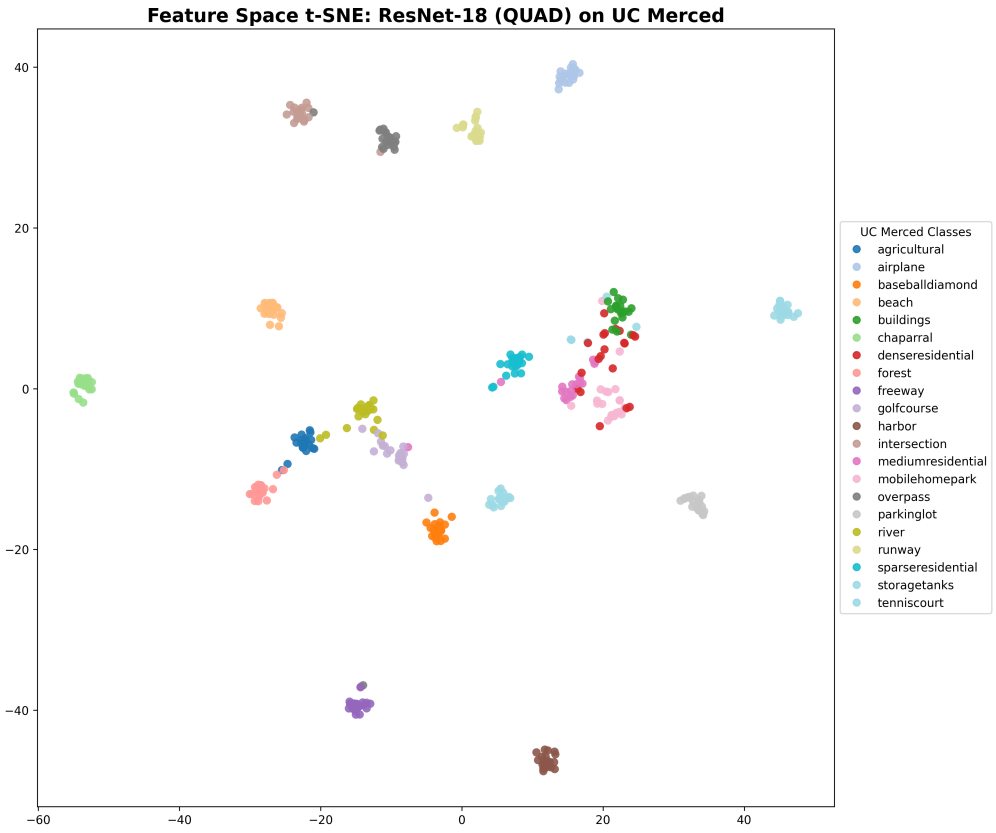

Figure 3: t-SNE feature space visualization of the test embeddings learned by our QuadContra framework on the UC Merced dataset ($\alpha = 0.05$, ResNet-18 backbone). The distinct, well-separated class clusters illustrate the effectiveness of our multi-relational objective in mitigating feature collapse under severe client heterogeneity.

On the UC Merced dataset, the QuadContra framework consistently surpasses all evaluated baseline methods. Under the most extreme data heterogeneity ($\alpha = 0.05$), our method achieves an accuracy of 88.23%, significantly outperforming the previous state-of-the-art RS-CCL, which achieves 86.00%[cite: 7]. This performance advantage holds across all heterogeneity levels, ultimately converging to 96.31% in the fully uniform IID setting.

Similarly, on the larger and intrinsically more challenging NWPU-RESISC45 dataset, the VGG16 model fortified by our contrapositive objective demonstrates superior resilience against client drift. At $\alpha = 0.05$, QuadContra reaches 86.06%, exceeding the RS-CCL baseline of 84.03%[cite: 8]. In the IID distribution, our framework achieves a peak accuracy of 93.30%.

Crucially, the consistently tighter convergence variance observed across both datasets verifies that the multi-relational geometric constraints of QuadContra effectively stabilize the VGG16 feature space. This empirically proves that our approach fundamentally resolves local dimensional collapse regardless of the underlying convolutional architecture.