Non-IID Data Impact on Multimodal Federated Model Alignment Efficiency

Assignee Research

May 31, 2026

Abstract

This report synthesises findings from 10 peer-reviewed papers addressing the following research question: What is the impact of varying degrees of non-IID data on the alignment efficiency of collaborative multimodal models in federated learning, as measured by structural similarity metrics (e.g., cosine. Federated learning learns from scattered data by fusing collaborative models from local nodes. However, due to chaotic information distribution, the model fusion may suffer from structural misalignment with regard to unmatched parameters. 9 claims were extracted from source literature; 0 were independently verified against retrieved documents. An automated multi-reviewer quality assessment produced a score of 3.1/10. This report is a machine-generated literature synthesis and does not constitute original research.

1 Introduction

This paper examines: Heterogeneous Federated Learning. Research question: What is the impact of varying degrees of non-IID data on the alignment efficiency of collaborative multimodal models in federated learning, as measured by structural similarity metrics (e.g., cosine similarity of feature embeddings) across clients?.

2 Methodology

Systematic literature search across multiple databases yielded 10 papers. Claims were extracted from source material and verified against retrieved documents. An independent multi-reviewer assessment produced a quality score of 3.1/10.

3 Results

10 papers retrieved. 9 claims extracted; 0 independently verified. Quality review score: 3.1/10.

4 Limitations

This report is a machine-generated literature synthesis and does not constitute original research. Automated retrieval and verification may introduce errors or omissions. Review scores reflect automated assessment, not human peer review. Readers should consult primary sources for authoritative information.

5 Extracted Claims

Claim	Verified	Confidence
The proposed framework achieves +3% to +5% accuracy improvement over FedMA under different local training epoch settings	×	0.03
The proposed method achieves +0.6% to +2.4% accuracy improvement over FedMA with similar or less local training budgets.	×	0.05
The proposed method achieves +3.0% and +4.0% accuracy improvement over FedAvg and FedProx, respectively, with the same t	×	0.03
The proposed framework achieves 89.96% accuracy with local epoch 1, 89.10% with local epoch 20, 89.28% with local epoch	×	0.02
FedMA achieves 87.53% accuracy, FedAvg achieves 86.29% accuracy, and FedProx achieves 85.32% accuracy.	×	0.01
The proposed framework achieves 88.26% accuracy with Group Normalization (GN), while FedAvg achieves 85.46% with Batch N	×	0.04
The proposed framework achieves 68.23% accuracy with 10 groups, 68.17% with 20 groups, and 67.23% with 100 groups, while	×	0.01
The proposed framework enforces structurally-aligned feature encoding and alleviates feature averaging conflicts among d	×	0.09
The proposed framework designs a feature-oriented regulation method (Ψ-Net) to identify local models' structural hierarc	×	0.12

References

- `http://arxiv.org/abs/2206.05507v1`

- http://arxiv.org/abs/2008.06767v2
- http://arxiv.org/abs/2506.09781v2