Non-IID Data Impact on Multimodal Model Inference Efficiency in Federated Learning

Assignee Research

May 31, 2026

Abstract

This report synthesises findings from 12 peer-reviewed papers addressing the following research question: How does varying degrees of non-IID data across federated clients influence the inference efficiency and throughput of multimodal models during distributed training. 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. 20 claims were extracted from source literature; 0 were independently verified against retrieved documents. An automated multi-reviewer quality assessment produced a score of 2.8/10. This report is a machine-generated literature synthesis and does not constitute original research.

1 Introduction

This paper examines: Heterogeneous Federated Learning. Research question: How does varying degrees of non-IID data across federated clients influence the inference efficiency and throughput of multimodal models during distributed training?.

2 Methodology

Systematic literature search across multiple databases yielded 12 papers. Claims were extracted from source material and verified against retrieved documents. An independent multi-reviewer assessment produced a quality score of 2.8/10.

3 Results

12 papers retrieved. 20 claims extracted; 0 independently verified. Quality review score: 2.8/10.

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 shows the best performance under all local epoch settings, improving FedMA by +3%$\sim$5% accuracy.	×	0.03
The proposed method achieves +0.6%$\sim$2.4% accuracy than 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 enforces structurally-aligned feature encoding and alleviates feature averaging conflicts in both	×	0.09
The proposed framework achieves 89.96% accuracy with local epoch 1.	×	0.03
The proposed framework achieves 89.10% accuracy with local epoch 20.	×	0.03
The proposed framework achieves 89.28% accuracy with local epoch 40.	×	0.03
The proposed framework achieves 88.10% accuracy with local epoch 100.	×	0.03
The proposed framework achieves 88.29% accuracy with local epoch 150.	×	0.03
FedMA achieves 87.53% accuracy.	×	0.00
FedAvg achieves 86.29% accuracy.	×	0.00
FedProx achieves 85.32% accuracy.	×	0.00
The proposed framework achieves 88.26% accuracy with Group Normalization (GN).	×	0.02
FedAvg achieves 85.46% accuracy with Batch Normalization (BN).	×	0.01
FedAvg achieves 83.34% accuracy with Group Normalization (GN).	×	0.00
FedAvg without normalization achieves 84.13% accuracy.	×	0.00
The proposed framework achieves 68.23% accuracy with 10 groups.	×	0.02
The proposed framework achieves 68.17% accuracy with 20 groups.	×	0.02
The proposed framework achieves 67.23% accuracy with 100 groups.	×	0.02
FedAvg achieves 65.32% accuracy.	×	0.00

References

- http://arxiv.org/abs/2508.14769v2
- http://arxiv.org/abs/2008.06767v2
- http://arxiv.org/abs/2405.07925v1