Scaling Synthetic Sample Diversity for Robust Multimodal Foundation Models

Assignee Research

June 9, 2026

Abstract

This report synthesises findings from 15 peer-reviewed papers addressing the following research question: What is the impact of scaling the diversity of structurally consistent synthetic samples on the robustness of multimodal foundation models evaluated by zero-shot transfer accuracy on out-of-domain. 16 claims were extracted from source literature; 0 were independently verified against retrieved documents. An automated multi-reviewer quality assessment produced a score of 3.2/10. This report is a machine-generated literature synthesis and does not constitute original research.

1 Introduction

This paper examines: Synthetic ALS-EEG Data Augmentation for ALS Diagnosis Using Conditional WGAN with Weight Clipping. Research question: What is the impact of scaling the diversity of structurally consistent synthetic samples on the robustness of multimodal foundation models evaluated by zero-shot transfer accuracy on out-of-domain datasets like LVIS or Object365?.

2 Methodology

Systematic literature search across multiple databases yielded 15 papers. Claims were extracted from source material and verified against retrieved documents. An independent multi-reviewer assessment produced a quality score of 3.2/10.

3 Results

15 papers retrieved. 16 claims extracted; 0 independently verified. Quality review score: 3.2/10.

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 study used the EEGET-ALS Dataset containing raw 32-channel EEG recordings sampled at 256 Hz.	×	0.07
The dataset includes data from six ALS patients, with each contributing up to ten sessions over three to five months.	×	0.11
The dataset includes data from 170 healthy controls, with each contributing one session.	×	0.03
Every recording in the dataset comprises nine roughly two-minute blocks of motor imagery, actual movement, eye-tracking–	×	0.01
No filtering or artifact rejection was applied to the raw EEG data prior to processing.	×	0.05
ALS samples represent approximately 3.4% of the total dataset.	×	0.05
Raw EEG matrices were linearly rescaled into the [–1, 1] range using the formula (x - x.min()) / (x.max() - x.min()) * 2	×	0.03
The data scaling was applied file-wise rather than across the entire dataset.	×	0.02
Each segment's rows were treated as independent samples without per-channel splitting or z-score normalization.	×	0.01
Visual inspection indicated that synthetic ALS signals exhibit oscillatory patterns and amplitude dynamics within the ra	×	0.12
Clinicians and EEG experts consulted found the synthetic EEG segments indistinguishable from real ALS EEG segments in ma	×	0.12
The critic's loss settled into a narrow oscillation around a stable mean after an initial adjustment period.	×	0.02
The generator's loss steadily declined toward its asymptotic value during training.	×	0.03
Weight clipping was enforced to the range [-0.01, +0.01].	×	0.04
No sudden divergence or runaway oscillations were observed in the training loss curves.	×	0.05
A formal diversity analysis of the generated data was not performed.	×	0.03

References

- http://arxiv.org/abs/2506.16243v1
- http://arxiv.org/abs/2511.10212v1
- http://arxiv.org/abs/2408.07303v2