

ALF-LB Load Balancing and Robustness in Sparse MoE Models under Domain Shift

Assignee Research

June 6, 2026

Abstract

This report synthesises findings from 16 peer-reviewed papers addressing the following research question: How does the ALF-LB load balancing method affect the robustness of sparse MoE models to domain shift, as evaluated by accuracy drop on cross-domain code completion tasks (Python to JavaScript) from. 12 claims were extracted from source literature; 2 were independently verified against retrieved documents. An automated multi-reviewer quality assessment produced a score of 4.5/10. This report is a machine-generated literature synthesis and does not constitute original research.

1 Introduction

This paper examines: Learning Robust Stereo Matching in the Wild with Selective Mixture-of-Experts. Research question: How does the ALF-LB load balancing method affect the robustness of sparse MoE models to domain shift, as evaluated by accuracy drop on cross-domain code completion tasks (Python to JavaScript) from the MBPP benchmark?.

2 Methodology

Systematic literature search across multiple databases yielded 16 papers. Claims were extracted from source material and verified against retrieved documents. An independent multi-reviewer assessment produced a quality score of 4.5/10.

3 Results

16 papers retrieved. 12 claims extracted; 2 independently verified. Quality review score: 4.5/10.

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
SMoEStereo achieves state-of-the-art joint generalization on the ETH3D, KITTI, and Middlebury benchmarks using the same	×	0.09
SMoEStereo leverages pre-trained Vision Foundation Models for robust stereo matching with minimal cost.	✓	0.15
The proposed method integrates MoE LoRA with varying ranks and MoE Adapter layers with varying kernel sizes into Vision	✓	0.20
A lightweight decision network integrated within each MoE module dynamically selects relevant modules and deactivates le	×	0.11
Previous robust stereo matching methods typically rely on classic feature extractor backbones such as ResNet, UNet, and	×	0.08
SMoEStereo (DAMV2) achieves a Bad 3.0 score of 4.22 on KITTI 2012 and 4.86 on KITTI 2015.	×	0.01
SMoEStereo (DAMV2) achieves a Bad 2.0 score of 7.05 on the Middlebury benchmark.	×	0.02
SMoEStereo (DAMV2) achieves a Bad 1.0 score of 2.10 on the ETH3D benchmark.	×	0.02
The SMoE (Ours) method using DAMV2 ViT-Base has 6.81M training parameters and 2.86M test parameters.	×	0.06
Full-Finetuning of DAMV2 ViT-Small results in a Bad 3.0 score of 10.1 on KITTI 2012 and 14.3 on KITTI 2015.	×	0.02
SMoEStereo (DAMV2) outperforms SMoEStereo (DINOv2) on the KITTI 2012 benchmark with a Bad 3.0 score of 4.22 compared to	×	0.02
Former-CFNet (DINOv2) achieves an average score of 5.0 on the weather condition benchmarks listed in Table (p5).	×	0.01

References

- <http://arxiv.org/abs/2212.10264v1>
- <http://arxiv.org/abs/2507.04631v1>

- <http://arxiv.org/abs/2511.18488v2>