

OpenPangu-7B-MLA Robustness on MMSU Emotional Intent Detection Under Acoustic Noise

Assignee Research

June 8, 2026

Abstract

This report synthesises findings from 4 peer-reviewed papers addressing the following research question: Does OpenPangu-7B-MLA demonstrate improved robustness over prosody-exclusive models on MMSU emotional intent detection under varying acoustic noise conditions. 10 claims were extracted from source literature; 1 was independently verified against retrieved documents. An automated multi-reviewer quality assessment produced a score of 4.0/10. This report is a machine-generated literature synthesis and does not constitute original research.

1 Introduction

This paper examines: MMSU: A Massive Multi-task Spoken Language Understanding and Reasoning Benchmark. Research question: Does OpenPangu-7B-MLA demonstrate improved robustness over prosody-exclusive models on MMSU emotional intent detection under varying acoustic noise conditions?.

2 Methodology

Systematic literature search across multiple databases yielded 4 papers. Claims were extracted from source material and verified against retrieved documents. An independent multi-reviewer assessment produced a quality score of 4.0/10.

3 Results

4 papers retrieved. 10 claims extracted; 1 independently verified. Quality review score: 4.0/10.

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
MMSU encompasses a wider range of acoustic features spanning 47 distinct tasks.	×	0.08
MMSU increases reasoning complexity by requiring models to integrate paralinguistic, phonetic, and semantic information.	×	0.08
MMSU is the first benchmark to systematically incorporate linguistically grounded phenomena into spoken language underst	✓	0.22
MMSU includes 47 distinct tasks.	×	0.08
MMSU covers prosody, intonation, phonetics, rhetoric, syntactics, non-verbal, disfluency, audio, and perception.	×	0.10
MMSU is evaluated with 22 models, including 12 Speech-LLMs and 10 Omni Large Language Models (OmniLLMs).	×	0.10
Each instance in MMSU consists of an audio clip and a text prompt, with the model choosing one of four options (A–D).	×	0.02
Answer options in MMSU are randomly ordered and balanced across the dataset to avoid positional bias.	×	0.02
All models in MMSU are evaluated with the same optimized instruction-following prompts to ensure fairness and minimize p	×	0.01
Qwen2.5-Omni-7B incorrectly identified the intonation of the sentence as Fall-Rise Intonation instead of Failing Intonat	×	0.01

References

- <http://arxiv.org/abs/2506.04779v3>

- <http://arxiv.org/abs/0811.1099v1>
- <http://arxiv.org/abs/2509.11168v1>