

# Multi-Task Training on EchoMind Enhances OpenPangu-7B-MLA Emotional Alignment Across Acoustic Environments

Assignee Research

June 8, 2026

## Abstract

This report synthesises findings from 11 peer-reviewed papers addressing the following research question: To what extent does multi-task training on the EchoMind dataset improve OpenPangu-7B-MLA's emotional alignment scores across diverse acoustic environments relative to single-task fine-tuning. 10 claims were extracted from source literature; 1 was independently verified against retrieved documents. An automated multi-reviewer quality assessment produced a score of 4.3/10. This report is a machine-generated literature synthesis and does not constitute original research.

## 1 Introduction

This paper examines: Emotion Recognition in the Wild using Deep Neural Networks and Bayesian Classifiers. Research question: To what extent does multi-task training on the EchoMind dataset improve OpenPangu-7B-MLA's emotional alignment scores across diverse acoustic environments relative to single-task fine-tuning?.

## 2 Methodology

Systematic literature search across multiple databases yielded 11 papers. Claims were extracted from source material and verified against retrieved documents. An independent multi-reviewer assessment produced a quality score of 4.3/10.

## 3 Results

11 papers retrieved. 10 claims extracted; 1 independently verified. Quality review score: 4.3/10.

## 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 ensemble method achieved an accuracy of 67.75% on the validation set and 64.68% on the test set.	✓	0.16
The challenge baseline accuracy was 52.97% on the validation set and 53.62% on the test set.	×	0.12
The ensemble method outperformed the results obtained using the isolated modules (BN-only and CNN-only).	×	0.05
The integration of BN and CNN leads to the best performance in emotion recognition.	×	0.09
The system can associate unknown input features to correct labels, as indicated by darker cells on the main diagonal of	×	0.02
The methodology for emotion recognition uses both top-down and bottom-up approaches.	×	0.10
The bottom-up module involves three steps: face detection, face pre-processing, and CNN forward pass.	×	0.02
Face detection is performed using a commercial library.	×	0.04
The top-down module involves three steps: acquiring the scene descriptors, setting evidences in the BN, and estimating t	×	0.02
Scene descriptors are estimated using a commercial library, which can be replaced with a state-of-the-art algorithm.	×	0.04

## References

- <http://arxiv.org/abs/1709.03820v1>
- <http://arxiv.org/abs/2306.09372v1>

- <http://arxiv.org/abs/2108.12828v4>