Training Data Skew Effects on Confidence Uncertainty in Multimodal Models Under Domain Shift

Assignee Research

June 9, 2026

Abstract

This report synthesises findings from 15 peer-reviewed papers addressing the following research question: To what extent does training data distribution skew impact confidence uncertainty metrics in multimodal models under extreme domain shift conditions. 15 claims were extracted from source literature; 0 were independently verified against retrieved documents. An automated multi-reviewer quality assessment produced a score of 4.2/10. This report is a machine-generated literature synthesis and does not constitute original research.

1 Introduction

This paper examines: Tackling covariate shift with node-based Bayesian neural networks. Research question: To what extent does training data distribution skew impact confidence uncertainty metrics in multimodal models under extreme domain shift conditions?.

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 4.2/10.

3 Results

15 papers retrieved. 15 claims extracted; 0 independently verified. Quality review score: 4.2/10.

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 uses CIFAR and TINYIMAGENET datasets with corrupted test set versions provided by Hendrycks & Dietterich (2019	×	0.03
The architectures used in the experiments are VGG16, RESNET18, and PREACTRESNET18.	×	0.02
Three latent variable structures were tested: 'in' (input only), 'out' (output only), and 'both' (input and output).	×	0.05
The number of Gaussian components (K) in the variational posterior was set to 1, 2, or 4.	×	0.04
On in-distribution (ID) data, optimal performance is quite similar across different latent architectures.	×	0.04
On out-of-distribution (OOD) data, the optimal performance of using both input and output latent variables is similar to	×	0.09
Using only input latent variables produces slightly worse optimal performance on OOD data compared to using output or bo	×	0.08
The optimal gamma (γ) value is lower when the model uses both types of latent variables (z, s) compared to other configu	×	0.04
In experiments with RESNET18 on CIFAR10 using K=4 components and only latent output variables, the Negative Log-Likeliho	×	0.04
Higher γ prevents the model from memorizing random labels even when the majority of training labels are wrong (80% noise	×	0.06
Higher γ leads to improved performance on clean test sets because it prevents learning from noisy training labels.	×	0.08
The variational entropy H[q(Z)] decreases over time during training.	×	0.05
The model with higher entropy (M32) performs better than the model with lower entropy (M16) across all corruption levels	×	0.03
The parameter λ controls the severity of the generated implicit corruptions.	×	0.07
Explicit image corruptions g(x) are approximated by finding xc that minimizes the loss function L(xc) defined in Equatio	×	0.04

References

- http://arxiv.org/abs/2207.05796v1
- http://arxiv.org/abs/2206.02435v2
- http://arxiv.org/abs/2604.09529v1