

Parameter Scale and Performance Degradation in Low-Resolution Visual Logic Puzzles

Assignee Research

June 6, 2026

Abstract

This report synthesises findings from 11 peer-reviewed papers addressing the following research question: What is the correlation between model parameter scale and performance degradation on visual logic puzzles within the LogicVista dataset under low-resolution conditions. 17 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: Learn to Accumulate Evidence from All Training Samples: Theory and Practice. Research question: What is the correlation between model parameter scale and performance degradation on visual logic puzzles within the LogicVista dataset under low-resolution conditions?.

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. 17 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 study utilizes the MNIST, Cifar10, and Cifar100 datasets for standard supervised classification tasks.	×	0.02
The study utilizes the mini-ImageNet dataset for few-shot classification tasks.	×	0.03
The LeNet model is employed for experiments on the MNIST dataset.	×	0.02
The ResNet18 model is employed for experiments on the Cifar10 and Cifar100 datasets.	×	0.03
The ResNet12 model is employed for experiments on the mini-ImageNet dataset.	×	0.02
Strong regularization in existing evidential models pushes many training samples into or close to zero-evidence regions,	✓	0.19
The proposed RED model achieves a test accuracy of 99.10% \pm 0.02 on MNIST using the exp activation function.	×	0.03
The proposed RED model achieves a test accuracy of 76.43% \pm 0.21 on Cifar100 using the exp activation function.	×	0.03
Using the exp activation function, the RED model results in 0 samples with evidence $E \leq 0.01$ out of 50,000 tested sample	×	0.06
Using the exp activation function, the RED model results in 25,154 samples with evidence $E > 1.0$ out of 50,000 tested sa	×	0.06
With 10% of training data, the RED model using the exp activation function achieves 97.27% accuracy.	×	0.03
With 100% of training data, the RED model using the exp activation function achieves 76.43% accuracy.	×	0.03
The RED model using the Log loss (23) and exp activation achieves an accuracy of 76.43% \pm 0.2.	×	0.05
The RED model exhibits an In-Distribution Vacuity of 0.2729.	×	0.01
The RED model exhibits an Out-of-Distribution Vacuity of 0.7552 on the SVHN dataset.	×	0.02
The standard exp activation function yields an In-Distribution Vacuity of 0.3227.	×	0.02
The standard exp activation function yields an Out-of-Distribution Vacuity of 0.7081 on the SVHN dataset.	×	0.02

References

- <http://arxiv.org/abs/2402.04177v3>
- <http://arxiv.org/abs/2306.11113v2>
- <http://arxiv.org/abs/2512.05091v1>