Parameter-Efficient Fine-Tuning of CodeT5 for Cross-Distribution Vulnerability Detection

Assignee Research

June 9, 2026

Abstract

This report synthesises findings from 15 peer-reviewed papers addressing the following research question: Does fine-tuning CodeT5 on a subset of specialized CWE tasks reduce the performance gap between in-distribution and out-of-distribution vulnerability detection more effectively than general multitask. 17 claims were extracted from source literature; 2 were independently verified against retrieved documents. An automated multi-reviewer quality assessment produced a score of 4.7/10. This report is a machine-generated literature synthesis and does not constitute original research.

1 Introduction

This paper examines: How Does Fine-Tuning Impact Out-of-Distribution Detection for Vision-Language Models?. Research question: Does fine-tuning CodeT5 on a subset of specialized CWE tasks reduce the performance gap between in-distribution and out-of-distribution vulnerability detection more effectively than general multitask pretraining?.

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 4.7/10.

3 Results

15 papers retrieved. 17 claims extracted; 2 independently verified. Quality review score: 4.7/10.

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 Caltech-101, Stanford-Cars, Food-101, Oxford-Pets, and ImageNet-1k as in-distribution (ID) datasets.	×	0.05
For each ID dataset, the training set is constructed with 16 random samples per class by default.	×	0.04
The OOD test datasets include subsets of iNaturalist, Sun, Places, and Texture.	×	0.03
For ImageNet-1k as the ID dataset, ImageNet-O and OpenImage-O are used as additional OOD datasets.	×	0.04
The categories in each OOD dataset do not overlap with the categories in the corresponding ID dataset.	×	0.05
CLIP-B/16, which uses ViT-B/16 as the image encoder, is the default backbone for the main experiments.	×	0.02
ZOCLIP denotes pre-trained CLIP models without fine-tuning.	×	0.09
For CoOp and CoCoOp methods, the context length is set to 4.	×	0.02
For CoOp and CoCoOp methods, context vectors are initialized using the pre-trained word embeddings of the phrase 'a phot	×	0.03
CoCoOp is trained with a batch size of 1 for 10 epochs using SGD.	×	0.02
CoOp is trained for 100 epochs with a batch size of 32.	×	0.02
TipAdapterF is trained with a batch size of 256 using AdamW for 20 epochs.	×	0.02
Cosine scheduling is used for all training methods mentioned.	×	0.03
The evaluation metrics used are FPR95, AUROC, and ID classification accuracy (ID ACC).	×	0.05
Parameter-efficient fine-tuning generally improves OOD performance over the zero-shot counterpart when using a simple OO	✓	0.17
The MCM score consistently demonstrates the most promising performance compared to alternative OOD scores when coupled w	✓	0.18
Adapted CLIP models demonstrate nearly perfect OOD detection performance for ID datasets with fine-grained categories su	×	0.13

References

- http://arxiv.org/abs/2504.16584v1
- http://arxiv.org/abs/2406.05892v1
- http://arxiv.org/abs/2306.06048v3