Large-Batch Adversarial Training Effects on CodeT5 Accuracy in Code Completion

Assignee Research

June 9, 2026

Abstract

This report synthesises findings from 12 peer-reviewed papers addressing the following research question: What is the correlation between perturbation budget magnitude and accuracy degradation in CodeT5 models trained with large-batch adversarial examples on code completion tasks. 11 claims were extracted from source literature; 1 was independently verified against retrieved documents. An automated multi-reviewer quality assessment produced a score of 4.4/10. This report is a machine-generated literature synthesis and does not constitute original research.

1 Introduction

This paper examines: How do SGD hyperparameters in natural training affect adversarial robustness?. Research question: What is the correlation between perturbation budget magnitude and accuracy degradation in CodeT5 models trained with large-batch adversarial examples on code completion tasks?.

2 Methodology

Systematic literature search across multiple databases yielded 12 papers. Claims were extracted from source material and verified against retrieved documents. An independent multi-reviewer assessment produced a quality score of 4.4/10.

3 Results

12 papers retrieved. 11 claims extracted; 1 independently verified. Quality review score: 4.4/10.

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 StdCNN architecture consists of Conv(3,3,10), Conv(3,3,10), MP(2,2), Conv(3,3,20), Conv(3,3,20), MP(2,2), FC(50), Dr	×	0.01
The M1 architecture consists of Conv(5,5,20), Conv(5,5,20), FC(500), and SM(10).	×	0.01
The C1 architecture consists of Conv(5,5,64), MP(3,3), BN, Conv(5,5,64), MP(3,3), BN, FC(384), FC(192), and SM(10).	×	0.01
Architectures M1 and StdCNN were used for MNIST and Fashion MNIST experiments.	×	0.01
Models C1 and ResNet18 were used for CIFAR-10 experiments.	×	0.03
Input training data was augmented with random cropping and random horizontal flips by default.	×	0.03
PGD based attack results in the Appendix corresponding to FGSM attack plots were generated with a step size k = 40.	×	0.04
In the experiments reported in Section 3, momentum was set to zero.	×	0.01
Momentum was set to 0.9 for the benchmark experiments.	×	0.02
As the batch size increases, the test accuracy decreases.	✓	0.16
As the batch size increases, the associated FGSM test accuracy drops.	×	0.11

References

- http://arxiv.org/abs/2103.15670v3
- http://arxiv.org/abs/2303.15127v1

- http://arxiv.org/abs/2006.11604v1