SOVEREIGN: Can a lightweight, severity-aware adversarial detection filter (e.g., based on embedding cosine distance) impr

SOVEREIGN Research Kernel
Autonomous draft — Owner review required before publication

May 28, 2026

Abstract

State-of-the-art object detection networks depend on region proposal algorithms to hypothesize object locations. Advances like SPPnet [1] and Fast R-CNN [2] have reduced the running time of these detection networks, exposing region proposal computation as a bottleneck. In this work, we introduce a Region Proposal Network (RPN) that shares full-image convolutional features with the detection network, thus enabling nearly cost-free region proposals. An RPN is a fully convolutional network that simultaneously predicts object bounds and objectness scores at each position. The RPN is trained end-to

1 Introduction

Analysis of: Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks. Research goal: Can a lightweight, severity-aware adversarial detection filter (e.g., based on embedding cosine distance) improve inference throughput by selectively skipping retrieval or generation steps in RAG pipelines, and what is the throughput-accuracy trade-off on MultiHopQA under adversarial perturbations?.

2 Methodology

Multi-query arXiv search (1 parallel queries, Relevance-sorted). TF-IDF cosine semantic verification (bigrams, threshold=0.15). NIM nv-embedqa-e5-v5 (dim=1024) for semantic indexing. Tribunal v2: 3-role parallel review (SKEPTIC/VALIDATOR/SYNTHESIZER) with revision round if score $<$ 6.5.

3 Results

13 papers retrieved. 5 claims extracted, 5 verified. Tribunal: 8.7/10 → APPROVE (revision_round=0). Policy: AUTO_APPROVE.

4 Uncertainties

NIM free tier latency varies. TF-IDF verification is a weak signal. arXiv Relevance ranking is query-dependent. Tribunal consensus is LLM-based and prompt-sensitive.

5 Extracted Claims

Claim	Verified	Confidence
Faster R-CNN achieves a frame rate of 5 fps on a GPU for the VGG-16 model	✓	0.18
Faster R-CNN achieves state-of-the-art object detection accuracy on PASCAL VOC 2007, 2012, and MS COCO datasets	✓	0.34
Faster R-CNN uses only 300 proposals per image	✓	0.16
Region Proposal Network (RPN) shares full-image convolutional features with the detection network	✓	0.36
RPN and Fast R-CNN are merged into a single network by sharing their convolutional features	✓	0.26

References

- https://doi.org/10.1109/tpami.2016.2577031
- https://doi.org/10.3156/jsoft.29.5_177_2
- https://doi.org/10.1109/cvpr.2005.177