SOVEREIGN: How does LogiPart's hypothesis-first hierarchical partitioning compare to full-corpus LLM conditioning on per-

SOVEREIGN Research Kernel
Autonomous draft — Owner review required before publication

May 29, 2026

Abstract

This paper presents a new vision Transformer, called Swin Transformer, that capably serves as a general-purpose backbone for computer vision. Challenges in adapting Transformer from language to vision arise from differences between the two domains, such as large variations in the scale of visual entities and the high resolution of pixels in images compared to words in text. To address these differences, we propose a hierarchical Transformer whose representation is computed with Shifted windows. The shifted windowing scheme brings greater efficiency by limiting self-attention computation to non

1 Introduction

Analysis of: Swin Transformer: Hierarchical Vision Transformer using Shifted Windows. Research goal: How does LogiPart's hypothesis-first hierarchical partitioning compare to full-corpus LLM conditioning on per-token throughput and F1 score across GLUE subtasks under varying covariate shift magnitudes?.

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

10 papers retrieved. 9 claims extracted, 7 verified. Tribunal: 7.5/10 → 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
Swin Transformer achieves 87.3 top-1 accuracy on ImageNet-1K	✓	0.15
Swin Transformer achieves 58.7 box AP and 51.1 mask AP on COCO test-dev	✓	0.27
Swin Transformer achieves 53.5 mIoU on ADE20K val	✓	0.18
Swin Transformer achieves +2.7 box AP improvement over previous state-of-the-art on COCO	✓	0.16
Swin Transformer achieves +2.6 mask AP improvement over previous state-of-the-art on COCO	✓	0.16
Swin Transformer achieves +3.2 mIoU improvement over previous state-of-the-art on ADE20K	×	0.12
Swin Transformer has linear computational complexity with respective to image size	×	0.14
Swin Transformer uses shifted windowing scheme that limits self-attention computation to non-overlapping local windows	✓	0.23
Swin Transformer can serve as a general-purpose backbone for computer vision	✓	0.17

References

- https://doi.org/10.1109/cvpr.2014.81
- https://doi.org/10.1109/cvpr.2009.5206848
- https://doi.org/10.1109/iccv48922.2021.00986