

# SOVEREIGN: What is the impact of domain shift in image complexity (e.g., from synthetic to natural scenes) on the accurac

SOVEREIGN Research Kernel

Autonomous draft — Owner review required before publication

May 28, 2026

## Abstract

In Transformer architectures, tokens\textemdash discrete units derived from raw data\textemdash are formed by segmenting inputs into fixed-length chunks. Each token is then mapped to an embedding, enabling parallel attention computations while preserving the input's essential information. Due to the quadratic computational complexity of transformer self-attention mechanisms, token reduction has primarily been used as an efficiency strategy. This is especially true in single vision and language domains, where it helps balance computational costs, memory usage, and inference latency. Despite the

## 1 Introduction

Analysis of: Token Reduction Should Go Beyond Efficiency in Generative Models – From Vision, Language to Multimodality. Research goal: What is the impact of domain shift in image complexity (e.g., from synthetic to natural scenes) on the accuracy-latency trade-off of CAT-based token reduction versus fixed patch tokenization in multimodal video generation models, evaluated using FID and CLIP scores?.

## 2 Methodology

Multi-query arXiv search (4 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**

12 papers retrieved. 0 claims extracted, 0 verified. Tribunal: 6.2/10 → REVERSE (revision\_round=1). Policy: ESCALATE\_TO\_OWNER.

### **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.

### **References**

- <http://arxiv.org/abs/2304.03307v1>
- <http://arxiv.org/abs/2505.18227v3>
- <http://arxiv.org/abs/2410.02152v1>