

SOVEREIGN: Does the performance advantage of SMOES over hand-crafted or modality-agnostic routing on unseen chart types g

SOVEREIGN Research Kernel

Autonomous draft — Owner review required before publication

May 28, 2026

Abstract

Charts are high-density visualization carriers for complex data, serving as a crucial medium for information extraction and analysis. Automated chart understanding poses significant challenges to existing multimodal large language models (MLLMs) due to the need for precise and complex visual reasoning. Current step-by-step reasoning models primarily focus on text-based logical reasoning for chart understanding. However, they struggle to refine or correct their reasoning when errors stem from flawed visual understanding, as they lack the ability to leverage multimodal interaction for deeper com

1 Introduction

Analysis of: ChartSketcher: Reasoning with Multimodal Feedback and Reflection for Chart Understanding. Research goal: Does the performance advantage of SMOES over hand-crafted or modality-agnostic routing on unseen chart types generalize from ChartQA to other multimodal benchmarks (e.g., FigureQA, PlotQA) under similar distribution shifts?.

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

11 papers retrieved. 8 claims extracted, 0 verified. Tribunal: 5.3/10 → REVISE (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.

5 Extracted Claims

Claim	Verified	Confidence
ChartSketcher-72B achieves significant improvements on both chart-specific and general-purpose datasets compared to the	×	0.05
The Qwen2VL-72B model was used as the base model for initializing ChartSketcher-72B.	×	0.03
ChartSketcher-72B was trained using LoRA with a rank of 16, Alpha of 32, batch size of 64, and learning rate of 1e-4 dur	×	0.06
ChartSketcher-72B was evaluated on the ChartQA dataset and achieved improved performance over the baseline Qwen2VL-72B.	×	0.02
The maximum tree depth in MCTS is set to 8, the maximum number of child nodes is 3, and the simulation count limit is 15	×	0.01
ChartSketcher uses a sketching reasoning pipeline that includes a programmatic sketching library and visual feedback mec	×	0.13
ChartSketcher incorporates a reflection and draw-feedback loop during the reasoning process.	×	0.10
The model uses a voting mechanism where each question is evaluated 3 times, and correctness is determined by a majority	×	0.02

References

- <http://arxiv.org/abs/2602.09258v1>
- <http://arxiv.org/abs/2505.19076v1>

- <http://arxiv.org/abs/2510.04514v2>