Your First Reviewer Has Left 1 Review

Your first reviewer has left 1 review

Suho Hanssch
sschSoongsil University Suho Hansuhohan@soongsil.ac.kr

TODO: Keywords (comma-separated)

Abstract

TODO: Write the abstract here.

1 Introduction

TODO: Write the introduction here.

2 Related Work

We position this work at the intersection of retrieval-augmented generation (RAG), agentic AI (tool-using and multi-step LLM systems), and emerging "agentic" workflows for assisting scientific reviewing.

Retrieval-Augmented Generation (RAG). RAG combines parametric generation with non-parametric retrieval to improve factuality and coverage on knowledge-intensive tasks [**?**]. Subsequent work explores stronger fusion architectures and scaling retrieval+generation pipelines, including fusion-in-decoder style conditioning [**?**] and other retrieval-conditioned generators. In the context of paper understanding and review assistance, RAG is particularly relevant because it enables grounding claims in retrieved evidence (e.g., paper sections, cited works, or external corpora), which can reduce hallucinations and support verifiable summaries.

Agentic AI and tool-using LLM systems. Agentic AI systems extend LLMs beyond single-shot generation by enabling iterative reasoning, tool use, and multi-step planning/execution [**?**, **?**, **?**]. A representative line of work uses interleaved reasoning and acting (e.g., calling search, code, or structured tools) to solve complex tasks [**?**]. These agentic paradigms are a natural fit for reviewing, where an assistant may need to (i) retrieve evidence, (ii) verify claims, (iii) cross-check prior work, and (iv) synthesize structured feedback under constraints.

Agentic reviewer workflows (practitioner systems). Parallel to academic research, practitioner-facing systems have proposed structured "agentic" workflows for document analysis and reviewing. For example, the Stanford "Agentic Reviewer" system (PaperReview.ai) converts a paper PDF into a structured representation, retrieves and summarizes relevant prior work from arXiv via web search, and then generates a comprehensive review following a template; it also reports rubric-style sub-scores and studies agreement with public ICLR 2025 ratings [**?**]. We treat these workflows as complementary motivation: they demonstrate practical decomposition patterns (extraction $\rightarrow$ retrieval $\rightarrow$ synthesis $\rightarrow$ scoring) that can be operationalized and evaluated in a research setting.

3 Method

3.1 Problem Setup

TODO: Define the task, inputs/outputs, and evaluation target.

3.2 Proposed Approach

TODO: Describe your main algorithm/model.

3.3 Training and Implementation Details

TODO: Provide training procedure, hyperparameters, compute budget.

4 Results

4.1 Experimental Setup

TODO: Datasets, splits, metrics, baselines, and protocol.

4.2 Main Results

TODO: Present core results.

4.3 Ablation Studies

TODO: Add ablations if applicable.

4.4 Analysis

TODO: Error analysis, robustness, efficiency, etc.

5 Conclusion

TODO: Write the conclusion here.

6 Related Work (Additional Discussion)

Automatic peer review and scientific assistants. Recent systems use LLMs to support scientific writing and evaluation, including summarization, rubric-based scoring, and critique generation. Our emphasis differs in that we target evidence-grounded reviewing behaviors by combining retrieval (for grounding) with agentic control (for verification and iterative refinement), rather than treating review generation as a purely generative task.

How this work fits. Relative to standard RAG pipelines [**?**], our focus is not only on grounded generation but also on end-to-end reviewing actions (e.g., targeted retrieval, claim checking, and structured decision support). Relative to general agent frameworks [**?**, **?**], we specialize the agent loop and evaluation to the peer-review context.

A Appendix (Optional)

TODO: Add appendix content here if needed.