

Peer Review

Review of: "GeAR: Generation Augmented Retrieval"

Yan Zhang¹

1. Tianjin University, China

This work introduces Generation-Augmented Retrieval (GeAR), a novel framework that harmoniously unifies natural language understanding (NLU) and generation (NLG) to address critical limitations in conventional document retrieval systems. GeAR advances the field through several contributions: (1) A novel architecture integrating fusion and decoding modules to explicitly model fine-grained semantic relationships between queries and documents. (2) A systematic pipeline for synthesizing high-quality training data via LLM-driven simulation.

However, I have the following concerns:

1. What are the differences between the proposed GeAR and BLIP? Following the structure of Figure 2, it seems that there is no obvious difference. This architecture itself supports both comprehension tasks and generation tasks, so the innovativeness of the authors' proposed method needs further explanation.
2. In addition to the retrieval and localization performance, the computational complexity and the number of model parameters still need to be compared in detail.

Declarations

Potential competing interests: No potential competing interests to declare.