

Investigating Effective LLM-based In-Context Tool Use: What Matters and How to Improve

Yining ZHENG, Haiyang WEI, Jiahao LU, Linqi YIN, Yunke ZHANG, Chengguo XU,
Hetao CUI, Tianxiang SUN, Shuang CHEN, Xipeng QIU

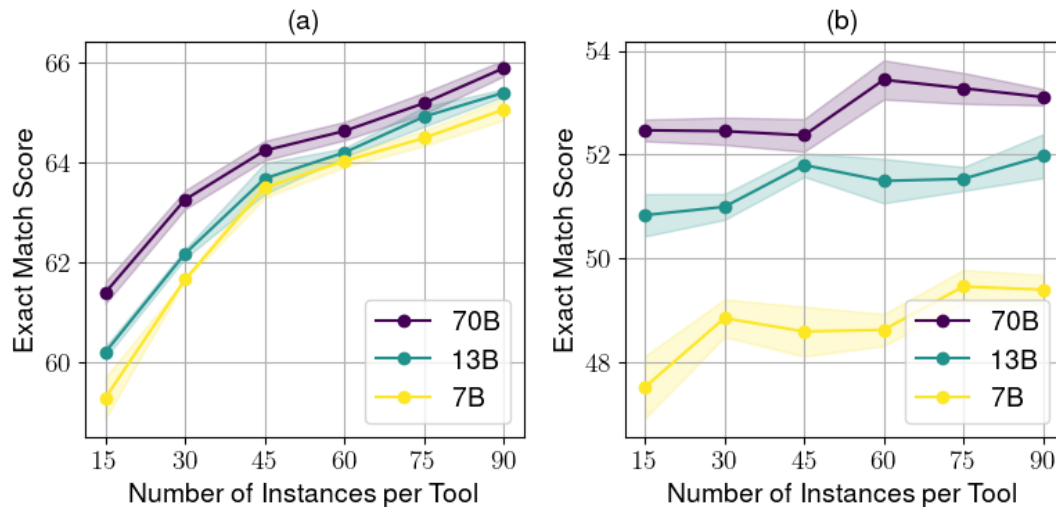
Frontiers of Computer Science, DOI: [10.1007/s11704-025-41365-6](https://doi.org/10.1007/s11704-025-41365-6)

Problems & Ideas

- Problems of Existing LLM-Based In-Context Tool Use:
 - LLMs can utilize tools after training, but methods for optimizing their performance remain unclear.
 - Existing research lacks a systematic investigation of the key factors influencing in-context tool use.
- Ideas:
 - Identify and analyze three key factors affecting in-context tool use:
 1. Number of tools in training data.
 2. Number of instances per tool in training data.
 3. Size of model parameters.
 - Construct RapidTools, a large-scale dataset for tool use research.
 - Conduct experimental evaluations to identify effective strategies for improving LLMs' in-context tool use capabilities.

Main Contributions

- Contributions:
 - Novel Dataset: RapidTools
 1. Over 192,393 instances, 2,124 target tools, and 4,699 candidate tools.
 2. Provides a robust benchmark for studying tool-use capabilities in LLMs.
 - Experimental Insights on Key Factors
 1. Increasing number of tools improves seen and unseen tool use.
 2. Increasing the number of training instances per tool consistently benefits performance on seen tools but leads to fluctuations on unseen tools.
 3. Increasing model parameter size significantly enhances unseen tool use.



Effect of increasing the number of training instances per tool on performance for models of different sizes on (a) seen tools and (b) unseen tools.