Doc-to-LoRA: Learning to Instantly Internalize Contexts

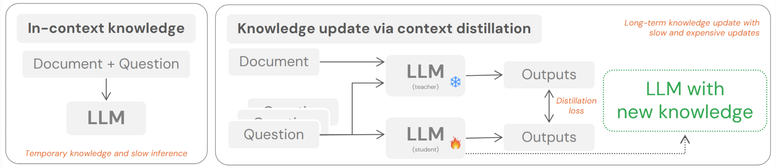
Rujikorn Charakorn, Edoardo Cetin, Shinnosuke Uesaka, Robert Lange

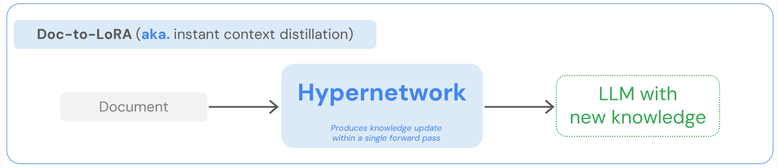
pub.sakana.ai/doc-to-lora/
Interactive demo and visualization

sakana.ai

ICML International Conference On Machine Learning

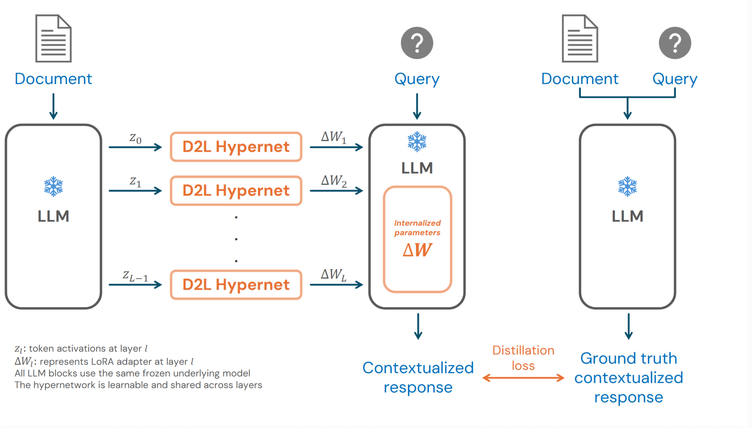
Motivation: Traditional knowledge update is slow; Doc-to-LoRA meta learns the knowledge update process



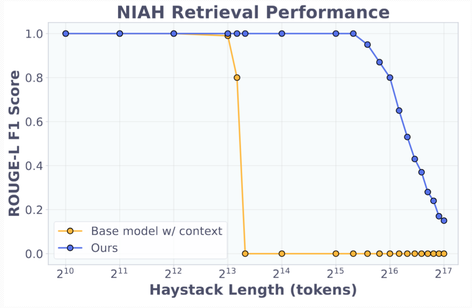


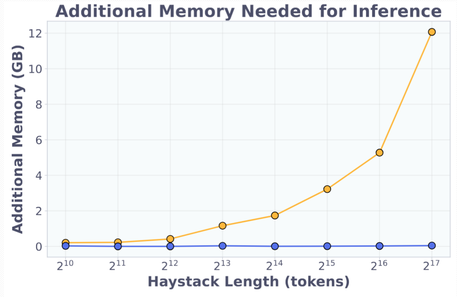
A hypernetwork that generates document-specific LoRAs from raw documents, updating the LLM parameters to answer related questions without the document in context

Training Overview: Meta-learning context distillation; Amortizing update cost upfront for cheap test-time updates



Internalizing needles beyond the context length





Base model: *Gemma-2-2b-it*

Training setup
-Context size 32-256 tokens
-Each sample is randomly "chunked" into 1 to 8 chunks
-LoRA from all chunks are concat'ed

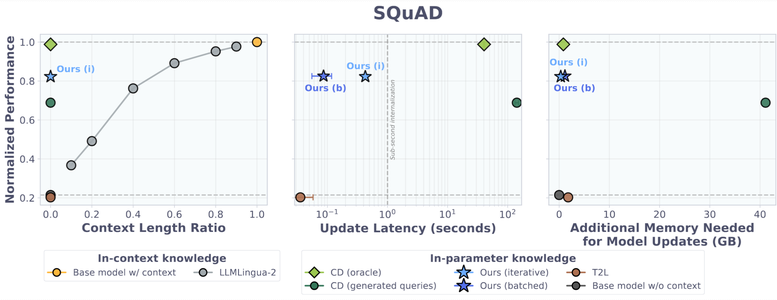
Eval setup
-Context size 1K to 128K
-Chunk size = 1024 tokens
-LoRA from all chunks are concat'ed

Doc-to-LoRA can effectively

(i) bypass the inherent context-length limitations of the base language model, and

(i) reduces the memory requirements for inference, especially under long context.

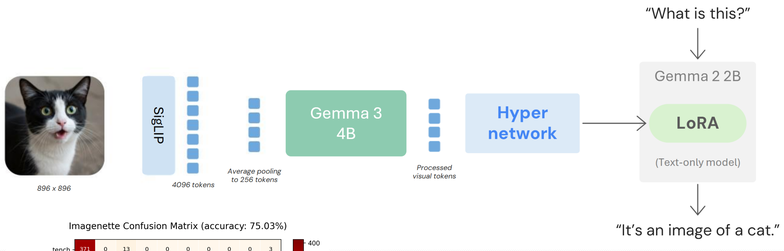
Real-world benchmark: Sub-second internalization; Answering document-based questions without doc in context

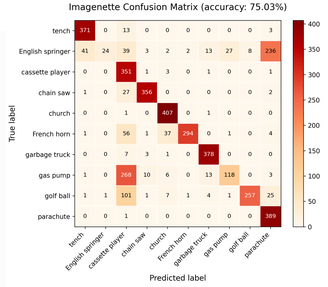


Training setup
-20M unique context, 5 QA pairs each
-Two stage training: no chunking (80%) → chunking (20%)
-Use web data + QA benchmark mixes

Eval setup
- Unseen questions and contexts

Zero-Shot Internalization of Visual Information





Training setup
- **Use VLM as the text encoder**
- 20M unique context, 5 QA pairs each (no images)
- Use web data + QA benchmark mixes

Eval setup
- **Use VLM as the image encoder**
- Map visual activations to LoRAs
- Image classification through internalized info

The target LLM can classify images with **75%** acc!