

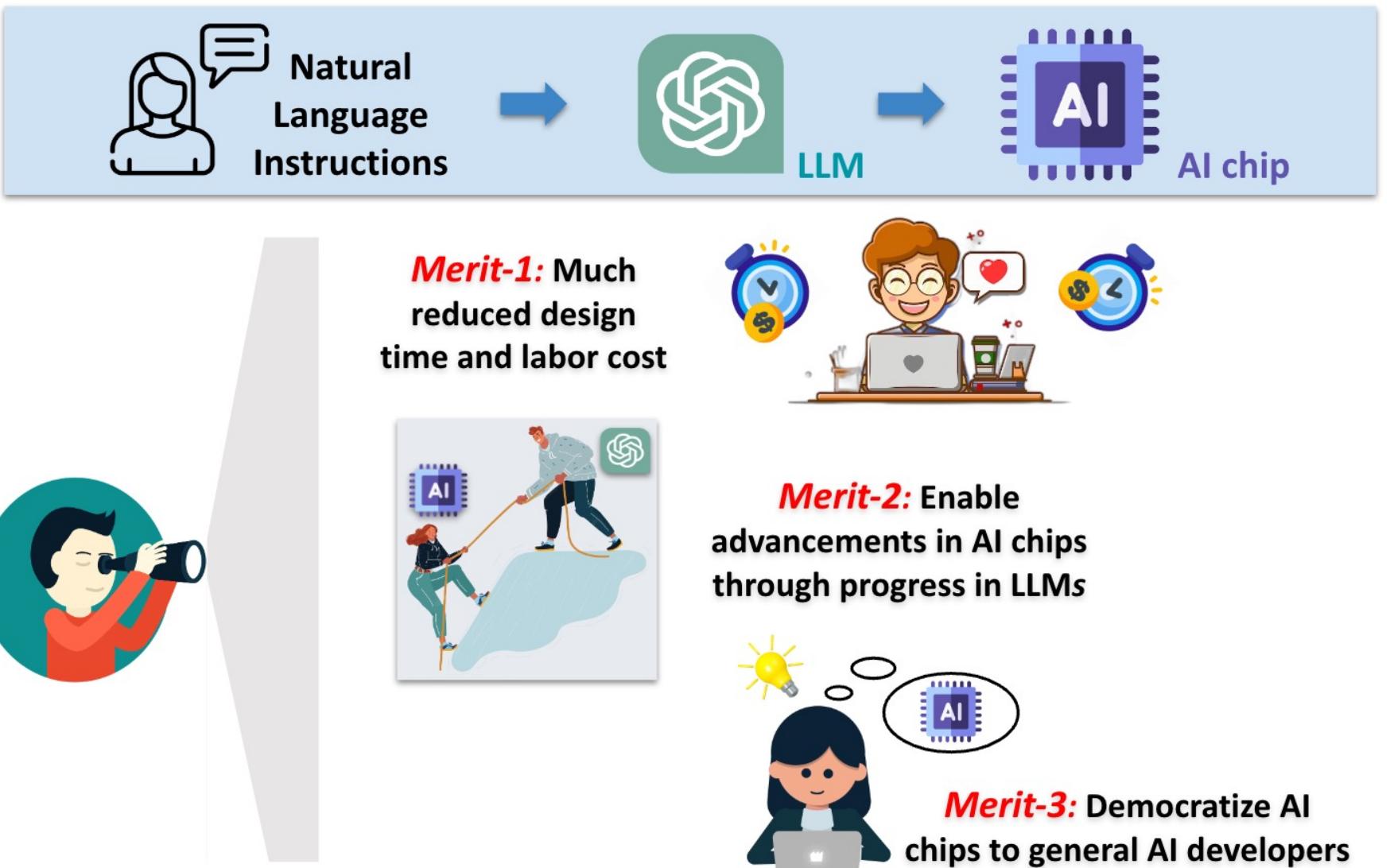
# GPT4AIGChip: Towards Next-Generation AI Accelerator Design Automation via Large Language Models

Yongan Zhang\*, Yonggan Fu\*, Zhongzhi Yu\*, Sixu Li, Zhifan Ye,  
 Chaojian Li, Cheng Wan, Yingyan (Celine) Lin  
 Georgia Institute of Technology



## GPT4AIGChip: Our Vision

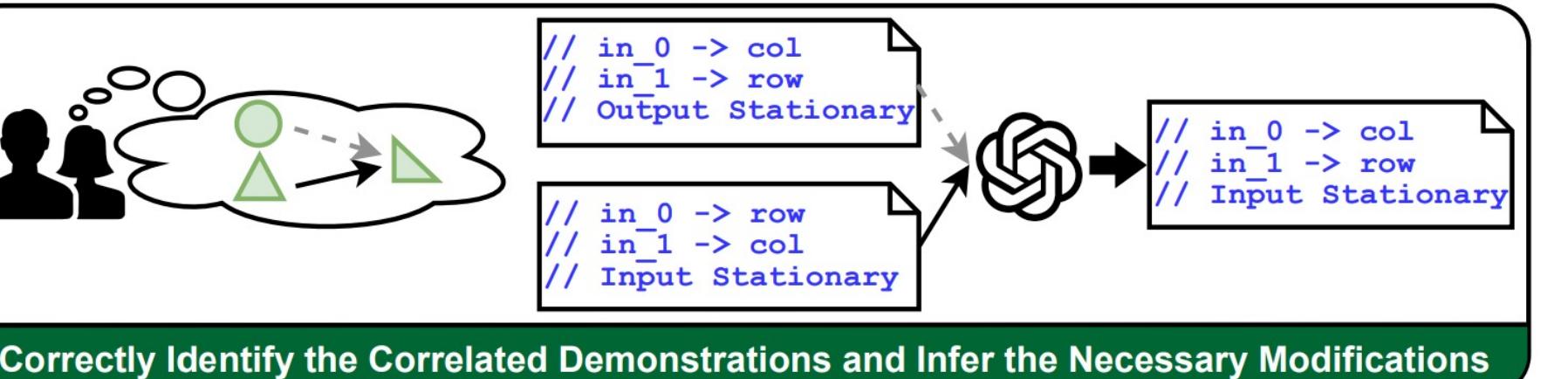
- ❖ Our vision: LLM-powered AI accelerator design using natural languages



## GPT4AIGChip: LLM Capabilities

- ❖ Identified capabilities of existing LLMs

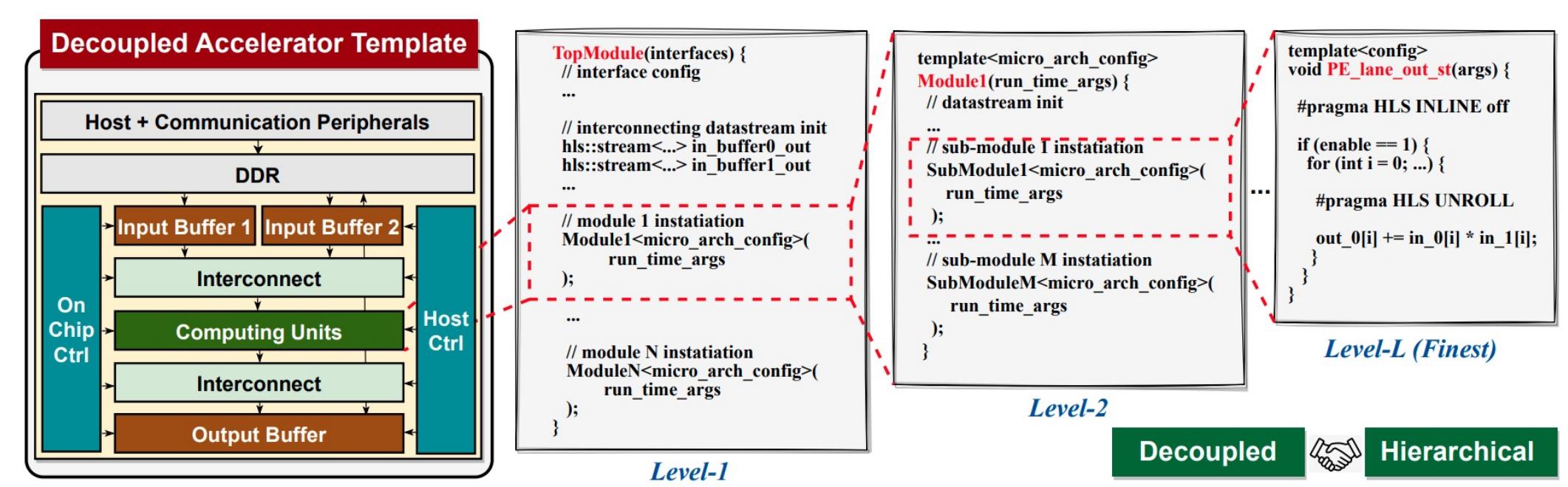
### Generalization and Logical Reasoning Capabilities



**Insights: Proper prompt engineering with demonstrations is essential to LLMs' success**

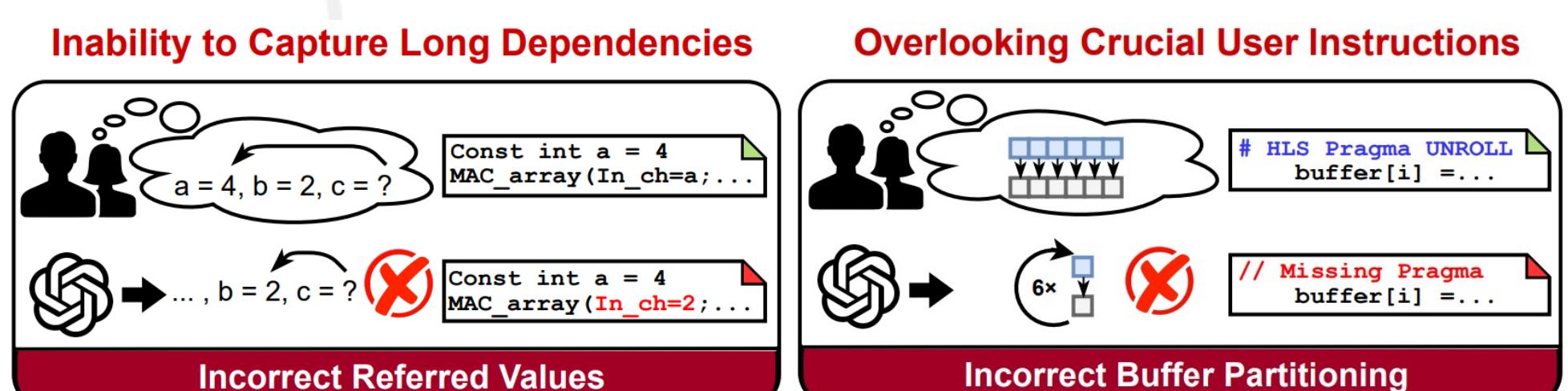
## GPT4AIGChip: Enabler 2

- ❖ Enabler 2: Decoupled Hardware Template



## GPT4AIGChip: LLM Limitations

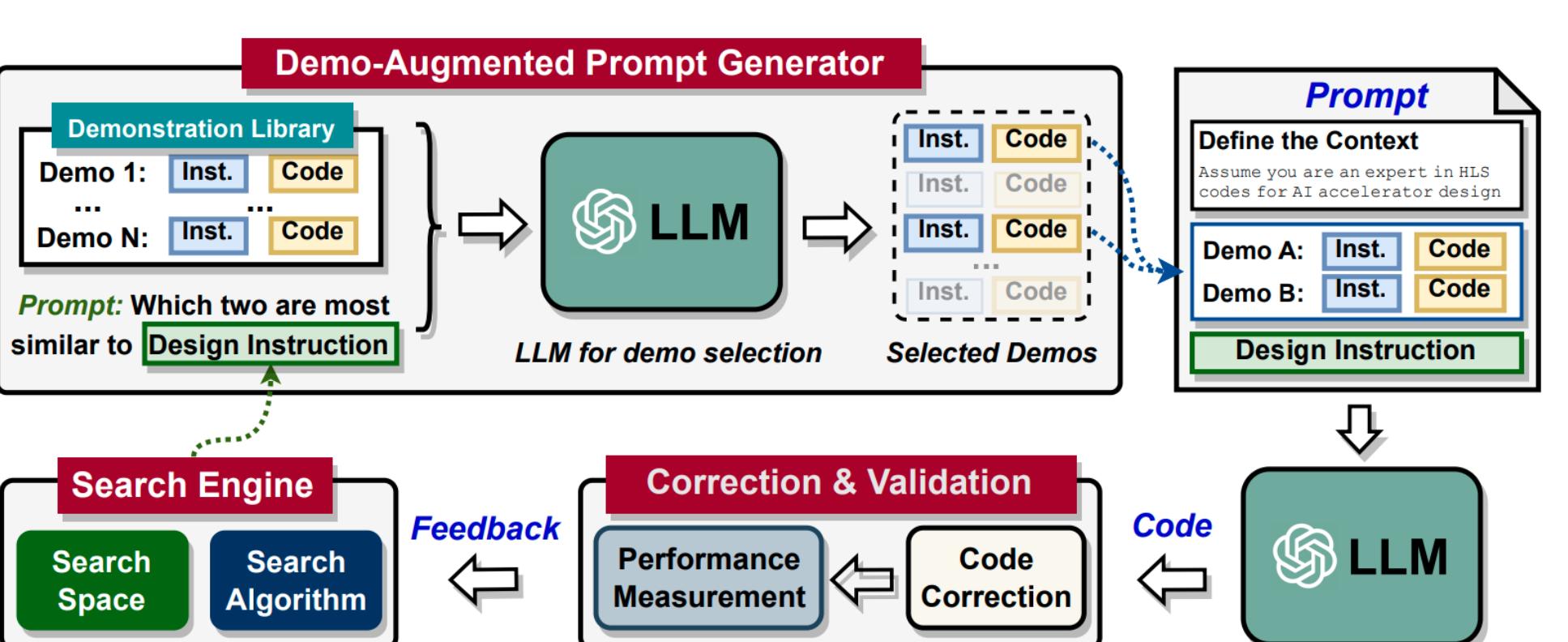
- ❖ Identified failures of existing LLMs



**Insights: Decoupling different hardware functionalities is crucial for design generation**

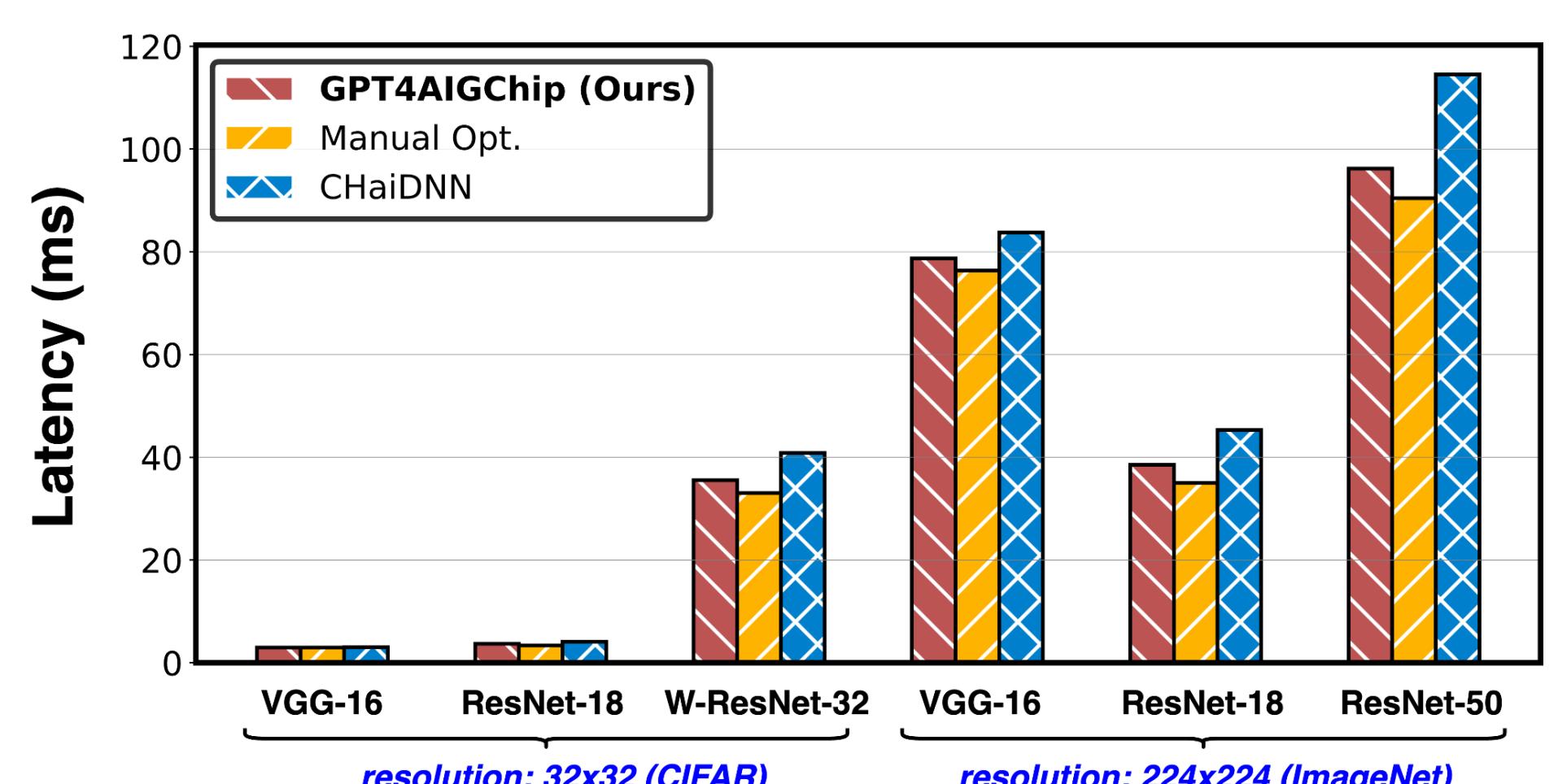
## GPT4AIGChip: Enabler 1

- ❖ Enabler 1: Demo-Augmented Prompt Generator



## GPT4AIGChip: Evaluation

- ❖ Against manual and automated accelerator designs across different workloads



Supported by NSF and CoCoSys, one of the seven centers in JUMP 2.0, a SRC program sponsored by DARPA.