Donghwan Chi

Research Statement

Motivation

My ultimate research goal is to achieve Artificial Super Intelligence (ASI), a form of machine intelligence surpassing human capabilities. I strongly believe in science as a powerful tool capable of fundamentally improving human society, and I see artificial intelligence (AI) as a pivotal meta-technology that accelerates this transformation. Just as recent developments in Large Language Models (LLMs) have significantly boosted developer productivity, I anticipate that AI advancements will increasingly accelerate progress across all fields. Currently, my primary research focus is on Multimodal Large Language Models (MLLMs), an essential step toward realizing Artificial General Intelligence (AGI) and ultimately ASI. I am particularly interested in extending and enhancing the exceptional reasoning capabilities of LLMs across various modalities.

My fascination with artificial intelligence began in high school when I witnessed AlphaGo's victory over Lee Sedol. That historic event inspired me to dedicate my life to AI research. Recent achievements, such as AlphaFold winning the Nobel Prize, have further solidified my conviction in this choice. Rather than passively awaiting the age of AI, I aspire to actively shape it as a researcher.

Previous Research Experience

During my Master's degree, I have conducted two significant research projects focusing on Multimodal Large Language Models.

The first project, titled "Slot-MLLM: Object-Centric Visual Tokenization for Multimodal LLM," addresses the limitations of existing image tokenization methods by proposing a novel approach for incorporating detailed, object-centric visual information into LLMs. By integrating the Slot Attention mechanism with the Q-Former structure, I developed discrete slot tokens capable of encapsulating precise, object-level information from images. This approach significantly improved image reconstruction quality compared to previous methods, enabling superior object-centric understanding and generation capabilities.

The second project, "Object-centric Self-improving Preference Optimization for Text-to-Image Generation," introduces a new Preference Optimization method to address the issue of MLLMs not accurately following textual instructions. By employing object-centric prompt perturbations and Visual Question Answering (VQA) evaluations, this method achieves high-quality image generation through a self-improving framework. Compared to traditional Preference Optimization methods, my approach provides meaningful differentiation at the detailed object level, significantly enhancing performance.

Through these research projects, I gained extensive practical experience. My involvement from the inception of a newly established laboratory allowed me to proactively tackle diverse challenges beyond research. In particular, I managed Linux servers and gained substantial experience building and maintaining Multi-GPU environments utilizing eight A100 GPUs and eight H100 GPUs. This experience provided me with valuable skills and confidence uncommon at the university level.

(+82) 010 6317 6021 • zheedong@korea.ac.kr

Research Goals

My research thus far has primarily focused on extending LLMs into multimodal environments. However, I plan to expand my future research to more fundamental and comprehensive domains. Specifically, I have outlined two primary research directions aimed at realizing AGI.

The first direction involves continuously enhancing the generalizability and multimodal capabilities of existing models. This includes developing methods to seamlessly integrate diverse modalities such as text, images, audio, and video into unified models capable of complex, real-world applications. My goal here is to address current limitations in model robustness, scalability, and flexibility to effectively handle multimodal data.

The second research direction focuses on advancing the reasoning capabilities of AI systems. I aim to develop sophisticated cognitive architectures capable of complex reasoning tasks such as causal inference, logical deduction, and long-term planning. Integrating reinforcement learning (RL) into these architectures will enable the creation of self-improving systems that autonomously adapt and optimize their own performance. Inspired by recent advancements like DeepMind's 'Reward is all you need,' I plan to explore novel reward structures and learning paradigms that facilitate more effective autonomous learning.

Furthermore, I intend to investigate how human feedback mechanisms, such as Reinforcement Learning with Human Feedback (RLHF), can be leveraged to further enhance AI alignment and performance. By effectively combining autonomous self-improvement techniques with targeted human feedback, I believe we can accelerate the development of intelligent systems that are not only powerful but also ethically aligned with human values.

In the long term, my research aims to develop AGI that surpasses human-level intelligence, capable of fundamentally addressing complex societal challenges such as climate change, healthcare, education, and resource management. Ultimately, I aspire to be a researcher who achieves significant academic accomplishments while delivering meaningful societal impacts through my work.