Sangkyu Lee

A oddqueue.github.io

Research Interest

Embodied Agents, Multimodal Models, AI Alignment, Long-context Modeling, Continual Learning

Education

Yonsei University

M.S. in Artificial Intelligence (Advisor: Youngjae Yu, CGPA: 4.16/4.30) Seoul, Korea

Yonsei University

Mar. 2021 – Feb. 2023 B.A. in Applied Statistics, B.S. in Computer Science and Engineering (CGPA: 4.06/4.30) Seoul, Korea

Mar. 2023 – Present

Mar. 2017 – Feb. 2021

Hanvang University

Major in Economics and Finance (Transferred, Includes a 2-year mandatory military service) Seoul, Korea

Research Experience

Mar. 2023 – Present **Graduate Research Assistant** Multimodal Intelligence Research Lab, Yonsei University (Advisor: Youngiae Yu) Seoul. Korea • Researching alignment of large language models and tokenized visual representation for embodied agents • Participated as a founding member and assisted with grant proposals for robotics foundation models **Research Scientist Intern** Sep. 2024 – Present Language Lab, LG AI Research (Mentors: Honglak Lee, Hosung Song, Janghoon Han) Seoul. Korea • Researching adaptive length tokenization of visual modality for efficient autoregressive representations • Investigating adaptive control of KL constraint during training steps on direct alignment algorithms **Technical Lead** Jan. 2024 – May 2024 World Model for Robotics and Vehicle Control Team, maumAI Seongnam, Korea • Led the development project of a world model for robotics agents and autonomous vehicles • Researched visual tokenizers aligned with text representations for fast adaptation of language models **Research Scientist Intern** Jul. 2023 – Jan. 2024 Foundation Research Team, Naver Cloud (Mentors: Kang Min Yoo, Sungdong Kim) Seongnam, Korea

- Participated in the alignment of HyperCLOVA X, a large language model specialized in Korean
- Researched self-alignment of language models through a self-assessment process by LLM-as-a-judge

Undergraduate Research Assistant Jul. 2022 – Aug. 2022 Big Data Analytics and Learning Lab, Yonsei University (Advisor: Noseong Park) Seoul, Korea

- Researched Neural Operators for the vision models to achieve resolution-free generative models
- Investigated Stratonovich SDE for accelerating the reverse-time process of diffusion models

PUBLICATIONS

KL Penalty Control via Perturbation for Direct Preference Optimization Sangkyu Lee, Janghoon Han, Hosung Song, Stanley Jungkyu Choi, Honglak Lee, Youngjae Yu	$\frac{\text{Under Review}}{\iota}$
SEAL: Entangled White-box Watermarks on Low-Rank Adaptation Giyeong Oh, Saejin Kim, Woohyun Cho, <u>Sangkyu Lee</u> , Jiwan Chung, Dokyung Song, Youngje	<u>Under Review</u> ae Yu
MASS: Overcoming Language Bias in Image-Text Matching Jiwan Chung, Seungwon Lim, <u>Sangkyu Lee</u> , Youngjae Yu	AAAI 2025
HyperCLOVA X Techinal Report HyperCLOVA X Team	<u>arXiv</u>
Aligning Large Language Models by On-Policy Self-Judgment Sangkyu Lee, Sungdong Kim, Ashkan Yousefpour, Minjoon Seo, Kang Min Yoo, Youngjae Yu	<u>ACL 2024</u>
Honors and Awards	
Partial Scholarship , Yonsei University (Mar. 2023) Academic Honors , Yonsei University (Jan. 2022)	
ACADEMIC SERVICES	
Reviewer, ICLR (2025)	

TEACHING EXPERIENCE

Teaching Assistant, AAI5012: Advanced Video Processing (Sep. 2023 – Dec. 2023) **Lecture Instructor**, Samsung AI Specialist Course (Jul. 2023)

EXTRACURRICULAR ACTIVITIES

Student Society: Lead of Data Science Team, YBIGTA, Yonsei University (Mar. 2022 – Feb. 2023) Club Activity: Band Leader (Lead Vocalist), SRGB, Hanyang University (Mar. 2017 – Feb. 2019)

TECHNICAL SKILLS

Programming Languages: Python, C⁺⁺, Java, R, MATLAB
Frameworks: PyTorch, PyTorch Lightning
Libraries: Transformers, TRL, Accelerate, DeepSpeed, vLLM, Diffusers, scikit-learn, Optuna
Math Coursework: Real Analysis, Mathematical Statistics, Linear Algebra, Numerical Analysis
Language Proficiency: Korean (Native), English (Fluent), Japanese (Fluent)