

Research Interests

My research journey, which began with Vision-Language models, is now expanding to **World Models for Embodied AI**. Recently I've been focusing on **latent world models for decision-time planning**.

Education

POSTECH(Pohang University of Science and Technology)

M.S. IN COMPUTER SCIENCE AND ENGINEERING

- Supervised by **Prof. Suha Kwak in the Computer Vision Lab**

Pohang, S.Korea

Feb. 2025 - Present

University of Seoul

B.S. IN ARTIFICIAL INTELLIGENCE

- Total GPA of 4.09 / 4.5, Major GPA of 4.11 / 4.5 (**Cum Laude**)

Seoul, S.Korea

Mar. 2021 - Feb. 2025

Johannes Gutenberg University Mainz

EXCHANGE STUDENT MAJOR IN COMPUTER SCIENCE

- Studied Trustworthy AI

Mainz, Germany

Sep. 2023 - Feb. 2024

Publication

ACID: Action Consistency via Inverse Dynamics for Planning with World Models

Gawon Seo, Dongwon Kim, Suha Kwak - [PDF] [Project Page]

arXiv preprint

- Proposed a decision-time planning framework that enhances the realizability and computational efficiency of action-conditioned world models by introducing Inverse Dynamics-based cycle action consistency.

Planning in 8 Tokens: A Compact Discrete Tokenizer for Latent World Model

Dongwon Kim, Gawon Seo, Jinsung Lee, Minsu Cho, Suha Kwak - [PDF] [Project Page]

IEEE/CVF Conference on Computer Vision and Pattern Recognition(CVPR) 2026

- Compressed each visual observation into just 8 discrete tokens and build a latent world model on top, achieving 40x faster planning with competitive accuracy for real-time robotic control.

Planning in 16 Tokens: A Compact Discrete Tokenizer for Latent World Model

Dongwon Kim, Jinsung Lee, Gawon Seo, Minsu Cho, Suha Kwak

The Conference on Robot Learning Workshop (CoRLW) - Learning to Simulate Robot Worlds, 2025

Generating Accurate and Detailed Captions for High-Resolution Images

Hankyeol Lee, Gawon Seo, Kyounggyu Lee, Dogun Kim, Kyungwoo Song, Jiyoung Jung - [PDF]

arXiv preprint

- Enhanced VLM captioning accuracy via LLM-based object verification and region-specific refinement.

Enhancing Visual Classification using Comparative Descriptors

Hankyeol Lee, Gawon Seo, Wonseok Choi, Geunyoung Jung, Kyungwoo Song, Jiyoung Jung - [PDF]

IEEE/CVF Winter Conference on Applications of Computer Vision(WACV) 2025

- Proposed Comparative Descriptors to enhance image classification in VLMs by emphasizing subtle differences between similar classes.

Experience

Robotics and Computer Vision Lab

UNDERGRADUATE RESEARCHER (ADVISOR : PROF. JIYOUNG JUNG)

- Co-authored two papers focused on VLMs.

University of Seoul

Apr. 2024 - Jan. 2025

Power Bank Rental Service One Eleven

AI DEVELOPER

- Developed a recommendation system to prioritize store locations to install power bank rental machines, analyzing commercial and GIS data.

Seoul, S.Korea

Jun. 2023 - Jul. 2023

Awards & Honors _____

2025.02	CUM LAUDE , Awarded the CUM LAUDE for excellent academic records	University of Seoul
2024.11	AICOSS President's Award , Real-time gaze tracking for Barrier-free kiosk, 2024 CO-SHOW AIM Challenge	COSS
2024.09	2nd Place in 2D Semantic Segmentation , 2024 Autonomous Driving AI Algorithm Development Challenge	MDATASYNC
2023-2024	Dongyang E&P Honors , Awarded to Engineering students with high achievements	Dongyang E&P
2023.02	AICOSS President's Award , AICOSS Industry-University Linkage Deep Learning Competition	AICOSS
2022.03	Academic Awards , Awarded to students with high achievements throughout the semester	University of Seoul
2021-2022	Semester High Honors , Awarded to students with high achievements throughout the semester	University of Seoul