

YoonJin Oh

☎ 010.9312.3703 | ✉ dhdbsslw@korea.ac.kr | 🌐 dhdbsslw | 📍 Seoul, South Korea

EDUCATION

Korea University

Mar. 2025 – Aug. 2026 (Expected)

Master's degree in Artificial Intelligence, **AGI Lab** (Advisor: Prof. Sungwoong Kim)

- **GPA** : 4.36/4.5
- **Coursework** : Computer Vision, Advanced Natural Language Processing, Advanced Reinforcement Learning, Automatic Speech Recognition, Edge Neural Network, and More

Korea University

Mar. 2020 – Feb. 2025

Bachelor's degree in Business / Computer Science & Engineering (Double Major)

- **GPA** : 4.01/4.5
- **Coursework** : Artificial Intelligence, Machine Learning, Deep Learning, Advanced Machine Learning (Generative Models), Natural Language Processing, Introduction to Convex Optimization, Linear Algebra, and More

RESEARCH INTEREST

Multimodal LLM

- Unified Multimodal Understanding and Generation across Vision, Language, Speech, Audio, and Beyond

Human Alignment

- Human Alignment in Multimodal Models via Reinforcement Learning

PUBLICATIONS

OSPO: Object-Centric Self-Improving Preference Optimization for Text-to-Image Generation CVPR 2026

Yoonjin Oh, Yongjin Kim, Hyomin Kim, Donghwan Chi, Sungwoong Kim

- Proposed **OSPO**, a **self-improving** framework leveraging **object-centric** preference optimization to mitigate **fine-grained object-level hallucinations** in MLLM-based text-to-image generation.
- 📄 Paper : [\[Link\]](#) / 🔄 Code : [\[Link\]](#)

FiRe: Fine-grained Multimodal Reasoning for Enhanced Image Generation

Preprint

Yongjin Kim, **Yoonjin Oh**, Yerin Kim, Hyomin Kim, Jeeyoung Yun, Yu-Jung Heo, Min-Jun Kim, Sungwoong Kim

- Proposed **FiRe**, a **fine-grained multimodal reasoning** framework with **step-level RL**, enabling MLLMs to iteratively identify and correct object-level errors in text-to-image generation.
- 📄 Paper : [\[Link\]](#) / 🔄 Code : [\[Link\]](#)

Slot-MLLM: Object-Centric Visual Tokenization for Multimodal LLM

Preprint

Donghwan Chi*, Hyomin Kim*, **Yoonjin Oh**, Yongjin Kim, Donghoon Lee, Daejin Jo, Jongmin Kim, Junyeob Baek, Sungjin Ahn, Sungwoong Kim

- Proposed **Slot-MLLM**, integrating a **slot attention-based object-centric visual tokenizer (SlotTok)** into a unified MLLM framework for enhanced object-level image understanding and generation.
- 📄 Paper : [\[Link\]](#) / 🔄 Code : [\[Link\]](#)

PROJECTS

Unified Vision-Language-Speech LLM

Mar. 2026 – Present

- Developing an **Omnimodal LLM** that unifies understanding and generation across **vision, language, and speech modalities** within a single next-token prediction framework.
- Integrating **Janus-Pro**-based decoupled visual tokenization and **DualCodec**-based speech tokenization to enable fine-grained, human-aligned multimodal understanding and generation.

SOFTWARE COPYRIGHTS

Object-centric Self-improving Preference Optimization (OSPO) Framework

Korea University Research and Business Foundation. *Registration No. C-2025-040011*, Korea Copyright Commission, 2025.

HONORS AND AWARDS

SKYCC Hackathon: Pairing (2023) : 2nd place, Union of Computer Clubs from SNU, KU, and YU

Mirae Asset Financial Big Data Festival (2021) : 4th place, Mirae Asset Securities Co., Ltd

Korea University Smart Campus Datathon (2021) : 2nd place, Korea University Office of Digital Information

KUBS Future Talent Scholarship (2020-2024) : Full Tuition Scholarship (4 Years), Korea University Business School

SKILLS

Languages : Python, Pytorch, Pytorch Lightning, C++
Engineering

- Multi-GPU LLM Training : DDP, FSDP, DeepSpeed
- Multi-Node LLM Training : verl (for RL) , Ray
- Inference Acceleration : vLLM
- Model Serving: FastAPI

LANGUAGE PROFICIENCIES

English : TOEIC 880

Korean : First Language

EXTRACURRICULAR ACTIVITIES

Academic Society : Korea University Deep Learning Academy, *AIKU*

Academic Society : Korea University Software Development Academy, *DevKor*