Woohyun **Cho**

AI Researcher · Multimodal Large Language Models

192, Guwol-ro, Namdong-gu, Incheon, Republic of Korea, 21561, Rep. of KOREA

(+82) 10-6679-6036 | k106419@yonsei.ac.kr | www.k1064190.github.io | k1064190 | cwhy

Education

Yonsei University — *Seoul, S.Korea*

M.S. in Computer Science — *Sep. 2024 - Aug. 2026 (Expected)*

- Total GPA of 4.24/4.3 (Credits Taken: 30)
- Multimodal Laboratory under the supervision of Professor Youngjae Yu

Yonsei University — *Seoul, S.Korea*

B.S. in Computer Science — *Mar. 2018 - Aug. 2024*

- Total GPA of 3.79/4.3, Major GPA of 3.81/4.3 (Credits Taken: 141/102)

Technical Skills

AI Framework	PyTorch, Transformers, vLLM, Diffusers, Accelerate, PyTorch Lightning
Programming	Python, C++, Javascript, LaTeX
Languages	Korean, English
Others	Linux, Claude Code, OpenCode, Oh-my-opencode

Industry Experience

Developing Microphone Noise Reduction and Speech Enhancement model using AI — *Seoul, S.Korea*

Project Lead (LG Electronics) — *Mar. 2025 - Dec. 2025*

- Led the development of a low-latency, real-time speech enhancement model optimized for LG TV's WebOS environment.
- Optimized lightweight architectures (e.g., FastEnhancer, LiSenNet) using Knowledge Distillation and Quantization, and successfully ported models to ONNX Runtime and TFLite for seamless edge inference.
- Conducted extensive training and evaluation using open-source datasets such as the DNS Challenge, ensuring robust performance.
- Demonstrated superior performance in qualitative evaluations through a live demo, effectively suppressing background noise, howling, and echo.

Multimodal RAG System for Document Question Answering — *Seoul, S.Korea*

Assistant (Hyundai Motor) — *Sep. 2024 - Dec. 2024*

- Developed a robust parsing pipeline for Korean industrial PDFs, addressing font encoding issues by integrating multiple libraries (PyMuPDF, PDFMiner) and utilizing PaddleOCR and DocLayout-YOLO for layout-aware extraction.
- Benchmarked state-of-the-art models (GPT-4o, Qwen2.5-7B, Qwen2.5-72B, Qwen-VL 72B) accelerated with vLLM and FlashInfer, evaluating performance via ANLS and BERTScore to determine the optimal setup for Document VQA.

Research Experience

Revisiting Residual Connections: Orthogonal Updates for Stable and Efficient Deep Networks — *Seoul, S.Korea*

Second Author — *May. 2025 - Sep. 2025*

- Investigated the impact of orthogonal methods on the Vision Transformer (ViT-S) model across CIFAR-10, CIFAR-100, and Tiny-ImageNet datasets.
- Performed targeted ablation studies, which included orthogonally re-training a standard pre-trained model and dynamically introducing an orthogonal structure during the training process.
- Pre-trained the SmolLM model using Nanotron and the Mamba model using Transformers and Trainer on the Cosmopedia-v2 dataset.

MAVL: A Multilingual Audio-Video Lyrics Dataset for Animated Song Translation — *Seoul, S.Korea*

First Author — *Dec. 2024 - May. 2025*

- Developed a web-crawling pipeline to collect original English and dubbed versions of animation song lyrics and their corresponding videos from sources like Genius, LyricsTranslate, Last.fm, and YouTube.
- Built a custom web application to facilitate the manual, side-by-side alignment of lyrical lines and sections between the source (English) and dubbed versions, using the video for synchronization.
- Conducted benchmark evaluations on the collected dataset using state-of-the-art LLMs, including Gemini 2.0 Flash, GPT-4o, and Qwen-72B.
- Designed and proposed a novel framework, SylAVL-CoT, to enable multimodal lyrics translation using existing large-scale models.

SEAL: Entangled White-box Watermarks on Low-Rank Adaptation *Seoul, S.Korea*

Third Author *Mar. 2024, In Progress*

- Compared the performance of conventional LoRA and a SEAL watermarked LoRA by fine-tuning Stable Diffusion v1.5 with DreamBooth using Diffusers Library.
- Benchmarked the performance of standard LoRA against the SEAL watermarked LoRA by fine-tuning the Gemma-2B model from the Transformers library across multiple downstream tasks.

Publications

International

2025 **Neurips 2025**, Main Conference Poster as a **Second Author** *San Diego, U.S.A.*

2025 **Emnlp 2025**, Main Conference Poster as a **First Author** *Suzhou, China*

Extracurricular Activity

YAI (Yonsei AI Academic Club) *Seoul S.Korea*

Member *June. 2023 - Aug. 2024*

- Led a project on musical style transfer, utilizing Stable Diffusion-based AudioLDM to change the instrumentation of audio tracks.
- Fine-tuned the LADI-VTON model on the AIHUB dataset to create a Virtual Try-On system tailored for Korean body types.

Prometheus (AI Academic Club) *Seoul S.Korea*

Member *Mar. 2023 - Feb. 2024*

- Developed a detective story game utilizing the 'eastworld' open-source project, agent prompting techniques, the GPT-4o model, and presented a live demonstration at a public exhibition.
- Executed a project on image colorization by fine-tuning Stable Diffusion with ControlNet, and presented a live demonstration at a public exhibition.

Molgorithm (Yonsei Algorithm Academic Club) *Seoul, S.Korea*

Member *Jan. 2022 - Jun. 2022*

- Participated in weekly problem-solving sessions, tackling complex algorithmic challenges from platforms like Baekjoon Online Judge.
- Competed in multiple Problem Solving contests, including the Seoul ICPC Regional, applying advanced data structures and algorithms under pressure.

Others

English Proficiency

- 2024.09 - TOEFL 94/120
- 2024.02 - TOEIC 935/990

Certification

- 2020.09 - Industrial Engineer Information Processing