

Education

- **University of Maryland, College Park** College Park, MD, United States
M.S. in Computer Science (*Advisor: Ruohan Gao*) Sep 2025 - May 2027
- **Yonsei University** Seoul, Republic of Korea
B.S. in Computer Science - **Summa Cum Laude, Valedictorian** Feb 2018 - Feb 2025
- Absence for mandatory military service (Aug 2020 - May 2022)
- **Chungnam Science High School** Gongju, Chungnam, Republic of Korea
Gifted School for Mathematics and Science - **Early Honors Graduate** Feb 2016 - Jan 2018

Industry Experience

- **Microsoft** Redmond, WA, United States
Applied Scientist Intern (Sequence Modeling, AI Agents & Tooling) Jun 2026 - Aug 2026
 - Developed a unified transformer-based sequential modeling framework for Microsoft 365 Copilot Chat growth analytics, enabling next-best-action recommendation and touchpoint sequence optimization through a shared learned representation.
 - Improved next-best-action recommendation performance by **6.13% in Hits@1** and **3.72% in MRR**, while achieving **1.33× stronger rank correlation** for touchpoint sequence optimization versus prior approaches.
 - Productized the framework as a self-service AI plugin exposed through a local MCP server, using ONNX Runtime and standalone packaging, enabling LLM-driven natural-language interaction through GitHub Copilot Chat/CLI without local ML environment setup.
 - Collaborated cross-functionally with customer engagement, data science, and applied science teams to incorporate domain expertise and real-world operational constraints into model recommendations.
- **Miraflow AI** Kirkland, WA, United States
Founding Machine Learning Engineer (Video Generation, MLOps & Infrastructure) Mar 2024 - May 2025
 - Developed and deployed an end-to-end video generation service for lip-sync video synthesis from scratch, automating the collection and preprocessing of **3.7TB+** video-related data using optimized Bash scripting with multi-GPU and multiprocessing support on **8×A100 GPUs**, achieving over **90% CPU/GPU utilization**.
 - Improved distributed training stability and increased training speed by **1.5×**, while reducing video inference latency by **63.4%** through profiling, multi-GPU/process parallelism, computation scheduling, and low-latency operation replacements.
 - Built a production-grade Sanic ML server for API-driven training, inference, and job control, and designed Kubernetes orchestration that reduced estimated cloud costs by up to **62.5%** under typical workloads.
 - Built an internal evaluation process with curated edge cases to evaluate lip-sync quality, artifacts, storage efficiency, and inference time, enabling data-driven pipeline improvements.
 - Led research on efficient diffusion-based deepfake image detection (DistilDIRE) for TrueMedia.org, achieving **3.2× speedup** via knowledge distillation without sacrificing accuracy (ICMLW 2024).

Current Research

- **University of Maryland** (Multisensory Machine Intelligence Lab) College Park, MD, United States
Graduate Researcher (Embodied AI, Multimodal Learning)—*Advisor: Ruohan Gao* Jun 2025 - Present
 - Led *OdorSpaces*, a physics-based benchmark for visual-olfactory embodied navigation in 3D environments; developed a multimodal RL policy that outperformed baseline policies in source localization (CoRL 2026).
 - Built a simulation pipeline aligning **4.4TB+** of time-varying 3D odor fields with embodied environments, reducing storage by **61.4%** via compressed Zarr representations with optimized chunking and data layout for efficient temporal access during RL training.

Publications (* DENOTES EQUAL CONTRIBUTION)

- **OdorSpaces: Visual-Olfactory Embodied Navigation in 3D Environments**
Changyeon Lee, Kelin Yu, Zheng Wei, Shuna Ni, Ruohan Gao
CoRL 2026
- **Deepfake-Eval-2024: A Multi-Modal In-the-Wild Benchmark of Deepfakes Circulated in 2024**
Nuria Alina Chandra, Ryan Murtfeldt, Lin Qiu, Arnab Karmakar, Hannah Lee, Emmanuel Tanumihardja, Kevin Farhat, Ben Caffee, Sejin Paik, Changyeon Lee, Jongwook Choi, Aerin Kim, Oren Etzioni
CVPR 2026 Workshop (Authenticity & Provenance in the Age of Generative AI)

DistilDIRE: A Small, Fast, Cheap and Light Diffusion Synthesized Deepfake Detection

- Yewon Lim*, Changyeon Lee*, Aerin Kim, Oren Etzioni

ICML 2024 Workshop (Foundation Models in the Wild)

The Tug-Of-War Between Deepfake Generation and Detection

- Hannah Lee, Changyeon Lee, Kevin Farhat, Lin Qiu, Steve Geluso, Aerin Kim, Oren Etzioni

ICML 2024 Workshop (Data-centric Machine Learning Research)

Cost-Efficient and Effective Counter Unmanned Aerial System via Visual-Acoustic Sensing

- Changyeon Lee*, Dongju Yu*, Soyeon Cho*, Dane W. Hindsley, Halaevatu F. Patterson, Megan A. Clecak, Eric T. Matson

IEEE International Conference on Robotic Computing 2024

Towards Interpretable Controllability in Object-Centric Learning

- Jinwoo Kim*, Janghyuk Choi*, Jaehyun Kang, Changyeon Lee, Ho-Jin Choi, Seon Joo Kim

CVPR 2024 Workshop (Causal and Object-Centric Representations for Robotics)

Prior Research Experience

Purdue University (IITP K-SW Square)

West Lafayette, IN, United States

- Visiting Researcher (Real-Time Multimodal ML Systems)—Advisor: Eric Matson

Mar 2024 - Jun 2024

- Fully sponsored by the Ministry of Science and ICT of the Republic of Korea

- Led development of a cost-efficient, real-time audio-visual Counter-Unmanned Aerial System (CUAS), building low-latency audio classification and multi-output regression models for drone detection and tracking that achieved **90%+ accuracy** in real-world tests on standard laptop hardware (IEEE IRC 2024).
- Built a WebRTC-based pipeline for real-time audio/image streaming and browser-based model access with secure authentication.

Yonsei University (Computational Intelligence & Photography Lab)

Seoul, Republic of Korea

- Research Intern (Computer Vision & Video Understanding)—Advisor: Seon Joo Kim

Jul 2022 - Jun 2024

- Implemented a Transformer-based Space-Time Video Super-Resolution (STVSR) architecture enhanced with a MeMVit-inspired memory bank, achieving comparable performance to CNN-based benchmarks while reducing model parameters by **75.3%**.

Selected Projects

Scaling Transfer Q* to Multi-Agent Test-Time LLM Alignment

- github.com/changyeon2/Multi-Agent-Transfer-Q

- Developed a modular, multi-agent extension of the Transfer Q* test-time alignment framework, enabling agents to iteratively evaluate, critique, and refine LLM outputs for improved safety beyond single-agent decoding.
- Achieved up to a **12.96%** reward-score improvement over the single-agent baseline on BeaverTails and PKU-SafeRLHF, with an average **66.91%** pairwise safety win rate in an LLM-as-a-Judge evaluation.

Deep Reinforcement Learning for Custom Game Environment

- github.com/changyeon2/Custom-made-Game-RL

- Developed a custom Python-based game environment and implemented reinforcement learning agents using DQN and Soft Q-Learning.
- Discovered behavioral divergence between algorithms, with DQN relying on heuristic shortcuts while Soft Q-Learning learned a stable and generalizable obstacle-avoidance policy.

Skills

- **Programming Languages:** Python, C++, Java, SQL
- **ML Frameworks & Libraries:** PyTorch, TensorFlow, OpenCV, Scikit-learn, XGBoost, NumPy, Pandas, SciPy
- **Infrastructure & Tooling:** Docker, Kubernetes, Slurm, ONNX Runtime, MCP, Git, Linux, Bash, FFmpeg
- **Cloud & Data:** Azure ML, Azure Synapse/PySpark, AWS, GCP, Microsoft Data Stack (Cosmos, SCOPE, Nitro), PostgreSQL

Honors & Scholarships

- **Merit-Based Scholarship** - Yonsei University (Spring 2019, Spring 2023, Fall 2023, Fall 2024)
- **Merit-Based Scholarship** - Jung-Hun Foundation (Fall 2023, Fall 2024)
- **Highest Honors** - Yonsei University (Fall 2023)
- **Honors** - Yonsei University (Spring 2018, Fall 2018)