

## EXPERIENCE

---

- **Toss Bank** Seoul, South Korea  
*Data Scientist* *Jan 2026 - Present*
  - **On-premise AI Agent System:** Building an on-premise opencode agent system for internal AI-assisted development workflows.
  - **AI Authentication:** Developing AI-based authentication for face and ID card verification.
  - **Document AI:** Developing document information extraction systems for banking workflows.
- **SuperbAI** Seoul, South Korea  
*Machine Learning Engineer* *Sep 2021 - Jan 2026*
  - **[Lead] Vision Foundation Model:** Led development of a multi-modal vision foundation model for text/semantic-prompted object detection and segmentation; prepared a 1.1M-image in-house dataset and drove planning through API deployment and AWS Marketplace release.
  - **CVPR 2025 Competitions:** Won 2nd place in IOD and 4th place in FSOD challenges, validating the model on open evaluation benchmarks.
  - **[Lead] ML Training & Serving Platform:** Built AWS Batch based training infrastructure and optimized inference with TensorRT and Triton, improving throughput by 5x versus pure PyTorch serving; generated approximately 1,100 models and 60 endpoints in one year.
  - **Efficient Training:** Applied parameter- and memory-efficient techniques, reducing GPU memory usage by 65.6% and training time by 44.3% while preserving model performance.
  - **Interactive Segmentation:** Built browser, hybrid, and cloud-based segmentation tools using RepViT-SAM, FocalClick, SAM, and SAM2, increasing segmentation labeling speed by 25x.
- **Kakao Enterprise** Pangyo, South Korea  
*Machine Learning Engineer Intern* *Feb 2021 - Aug 2021*
  - **AutoML:** Fixed AutoGluon NeuralNetFastAI scaling issue and developed a Flask-based training/inference AutoML framework with a simple front-end.
- **Team ISLAND** Seoul, South Korea  
*Chief Technology Officer* *Mar 2019 - Oct 2020*
  - **Startup:** Raised approximately \$290K through government subsidies and seed funding; managed five developers across Android, Unity, and deep learning workstreams.
  - **Mobile ML:** Developed a lightweight 3D human pose estimation model for mobile devices and converted PyTorch models to TensorFlow Lite.

## OPEN SOURCE

---

- **Hugging Face Transformers - SAM2:** Led the addition of Segment Anything 2 support to Transformers, including model integration, image/video processors, conversion flow, documentation, and tests; credited in the official [SAM2 docs](#).
- **Hugging Face Transformers - Molmo2:** Opened Molmo2 support for Transformers and published [danelcsb/Molmo2-4B](#) on the Hugging Face Hub.
- **Hugging Face Contributions:** Implemented or improved RT-DETR, ViTPose, DETA training loss, and distributed training utilities in [Transformers](#).

## PAPERS

---

- **ZERO: Multi-modal Prompt-based Visual Grounding:** arXiv - Jul 6, 2025
- **MobileHumanPose: Toward real-time 3D human pose estimation in mobile devices:** IEEE/CVF CVPR Workshop - Jun 19, 2021

## PROJECTS

---

- **Stable Diffusion:** Generated personalized portraits and custom styles by fine-tuning diffusion weights with LoRA.
- **iOS Inspection App:** Built an in-house iOS data inspection tool for project managers and distributed it through TestFlight.

## EDUCATION

---

- **KAIST (Korea Advanced Institute of Science and Technology)** Daejeon, South Korea  
*Professional Master in Electrical Engineering; GPA: 3.55/4.3*  
*Feb. 2020 – Aug. 2021*
- **POSTECH (Pohang University of Science and Technology)** Pohang, South Korea  
*Bachelor in Electrical Engineering; GPA: 3.4/4.3*  
*Mar. 2015 – Feb. 2020*
- **UIUC (University of Illinois Urbana-Champaign)** Illinois, United States  
*Exchange Student; GPA: 3.6/4.0*  
*Jan. 2018 – Dec. 2018*

## PROGRAMMING SKILLS

---

- **Professional:** Python, PyTorch, AWS, Transformers  
Langfuse
- **Experienced:** TensorFlow, Swift, LangChain,