

# Donghyun Son

happydh1@snu.ac.kr / Google Scholar / Personal Website

## EDUCATION

University of Texas at Austin, Exchange Student, Computer Science

Sep 2025 — Present

Seoul National University, B.S. in Computer Science

March 2018 — Feb 2026 (expected)

Three-year leave of absence for mandatory military service, served as an industrial technical personnel

- **GPA:** 4.04/4.3 (4.11/4.3 in major)
- **Relevant Coursework:** Scalable High-Performance Computing (Graduate Course; 1st out of 100+ students in CUDA competition); Natural Language Processing (Graduate Course); Computer Vision; Mathematical Statistics; Linear Algebra;

## EXPERIENCE

Undergraduate Researcher

August 2025 — Present

UTNS Research Group @ UT Austin

Austin, Texas

Supervised by Prof. *Aditya Akella*

- Researching agentic systems for efficient multi-agent coordination and resource management

Undergraduate Research Intern

September 2024 — August 2025

CMALab @ SNU

Seoul, South Korea

Supervised by Prof. *Sungjoo Yoo*

- Investigated a calibration-free vector quantization method for KV cache, incorporating a novel normalization algorithm and efficient CUDA kernels for 1–2 bit inference (NeurIPS25).

Machine Learning Engineer, Moderation ML Team

March 2022 — July 2023

Hyperconnect (Acquired by Match Group)

Seoul, South Korea

- Developed an ML-based system for moderating images and audio across Match Group brands (e.g., Tinder, Hinge)
- Designed an efficient human-in-the-loop system for content moderation, managing AI flywheel
- Devised a threshold optimization algorithm based on GD to optimize the tradeoff between reliability and cost (WSDM23, Oral)
- Proposed a training algorithm for ML models to generalize effectively across different services and domains (OOD-CV@ICCV23)

Research Engineer, R&D Team

August 2020 — March 2022

VisualCamp

Seoul, South Korea

- Designed and implemented modern C++ based multi-threaded ML inference pipeline for gaze estimation SDK
- Built a robust training procedure to train an appearance-based gaze estimation model and improved accuracy by 30%
- Proposed a transformer-based calibration algorithm that predicts user-specific latent vectors using users' samples

## PUBLICATIONS

**NSNQuant: A Double Normalization Approach for Calibration-Free Low-Bit Vector Quantization of KV Cache**

Donghyun Son, Euntae Choi, Sungjoo Yoo

Advances in Neural Information Processing Systems (NeurIPS 2025, to appear)

**In-Context Learning with Noisy Labels**

Junyong Kang, Donghyun Son, Hwanjun Song, Buru Chang

Preprint

**Gradient Estimation for Unseen Domain Risk Minimization with Pre-Trained Models**

Byounggyu Lew<sup>†</sup>, Donghyun Son<sup>†</sup>, Buru Chang

Workshop and Challenges for Out-of-Distribution Generalization in Computer Vision @ ICCV2023 (OOD-CV@ICCV23)

**Looking to Personalize Gaze Estimation Using Transformers**

Seung Hoon Choi, Donghyun Son, Yunjong Ha, Seonghun Hong, Taejung Park

Journal of Computing Science and Engineering (JCSE), Vol. 17, No. 2, pp.41-50

**Reliable Decision from Multiple Subtasks through Threshold Optimization: Content Moderation in the Wild**

Donghyun Son<sup>†</sup>, Byounggyu Lew<sup>†</sup>, Kwanghee Choi<sup>†</sup>, Yongsu Baek, Seungwoo Choi, Beomjun Shin, Sungjoo Ha, Buru Chang

ACM International Conference on Web Search and Data Mining (WSDM 2023, Oral)

## PRESENTATIONS

---

- **KV Cache Compression for Long Context Inference**, Weekly Seminar @ Deepest S16 (January 2025)
- **Efficient Algorithms for LLM Inference**, Weekly Seminar @ Deepest S15 (July 2024)

## HONORS & AWARDS

---

### Scholarships

- **Samsung Software Membership**, from Samsung Research (November 2021 — Present)
- **Korea-U.S. STEM Student Exchange Scholarship** (\$9000), from Ministry of Trade, Industry and Energy (2025.09)
- **Full Tuition Academic Scholarship** (merit-based, \$2300), from Seoul National University (2020-1, 2023-2, 2024-1, 2024-2)

### Programming Contests

- **2<sup>nd</sup> award**, at Union of Clubs Programming Contest (UCPC) 2025
- **5<sup>th</sup> award**, at Samsung Collegiate Programming Contest (SCPC) 2021
- **5<sup>th</sup> place**, at ICPC NERC Huawei Challenge 2020: Cloud Scheduling Challenge
- **1<sup>st</sup> place**, at Seoul National University Programming Contest (SNUPC) 2019, div. 2

### Others

- **Bronze Medal**, at 42nd Undergraduate Mathematics Contest div.1 (for mathematical majors, by Korean Mathematical Society)
- **Microsoft Azure Champ Prize : Hack For Social Good**, at TartanHacks 2021 in CMU
- **Best Paper Award**, at HCLT 2020
- **Silver Prize**, at The International Mathematics Tournament of the Towns (ToTT) 2015

## ADDITIONAL INFORMATION

---

**Interests:** Table Tennis, Math Puzzles, Mind Sports (4 Dan in Go)

**Technical Skills:** C++, Python, Pytorch, Tensorflow, Jax, CUDA, OpenCL, MPI, OpenMP

**Problem Solving:** Codeforces (handle: **diordhd**, rating: 2200+) / BOJ (handle: **dhdroid**)

**TOEFL iBT:** 108 (R29/L29/S22/W28)