Sanghyun Son, PhD

Contact Information

Affiliation:	Samsung Electronics
Address:	130 Samsung-ro, Yeongtong-gu, Suwon-si, Gyeonggi-do, 16678, Korea
Email:	sonsang35(at)gmail.com
Github/Homepage:	https://github.com/sanghyun-son, https://sanghyun-son.github.io
Google scholar:	link

Personal Statement

I received my Ph.D. in Electrical and Computer Engineering from Seoul National University, where my research focused on image restoration and super-resolution. My academic work includes the widely cited EDSR paper, with over 8,000 citations and 2,500 stars, reflecting both scholarly and community impact.

At Samsung Electronics, I build production-ready systems for manufacturing process automation and model safety. I have worked closely with cross-functional teams to deploy fine-tuned LLMs to over 10K internal engineers, improving their daily workflows. I also led a design of scalable training pipelines for 70B models across 128+ GPUs and an orchestration platform that doubled GPU utilization. My work spans ML system architecture and resource optimization, with a strong focus on practical deployment and real-world applicability.

I have a strong interest in open-source AI frameworks and scalable system design. All of my research code is publicly available, and I have made several contributions to deep learning community such as the MLPerf benchmark and international workshop collaborations. I value the seamless integration of technology into everyday life and aim to contribute to building systems that people can trust and rely on.

Work Experience

Staff Research Engineer at Samsung Electronics (AI Development, AI Center)

LRM and Agentic Workflow for Manufacturing Process Automation — Dec. 2024 – Present

- Implemented and debugged a distributed RL system for training in-house LRMs. Orchestrated over 128 GPUs using ray, utilized FSDP and vLLM for 70B model optimization and rollout, respectively.

LLM-based Safeguard System for In-house Models. — Aug. 2024 – Dec. 2024

- Developed and released (for >10K internal users) a safeguard model for in-house LLM infrastructure. Collected and synthesized harmful Korean data (∼50K), performed PEFT/full fine-tuning of 8B models using DeepSpeed/transformers.
- Organized a development team of ∼10 researchers and engineers for in-house model orchestration and deployment. Developed and managed the in-house API server using FastAPI.

Domain-Specific LLMs for Manufacturing Process Automation. — Apr. 2024 – Jan. 2025

- Developed a fine-tuning strategy using DeepSpeed and prompting pipeline with DSPy for a domain-specific LLM trained on ∼10K in-house instruction data.
- Achieved 8× smaller model size while outperforming larger general-purpose models.
- Deployed to >10K internal engineers and integrated into a daily-updated system.
- Led a TF team of 7 researchers for multi-modal (audio/video/image/documents) data curation.
- Built an automatic LLM resource orchestrator system as a solo side project, improving GPU utilization efficiency by over 2×; now actively used every day by our team.

AI-based ISP Solution for Low-light Environments. — Sep. 2023 – Mar. 2024

- Led a team of 3 researchers and developed PyTorch-based low-light machine vision systems for automotive, based on real hardware ISP pipeline.

Industry & Academic Collaborations

Core Technical Contributor, MLPerf Mobile AI Benchmark: Super-Resolution Track
In conjunction with **MLCommons, Google, Qualcomm, Samsung** Apr. 2022 – Mar. 2023

- Released an official mobile super-resolution benchmark integrated into the MLPerf application.
- Drove the technical direction of a cross-company collaboration to design and develop the benchmark.
- Proposed a mobile-friendly EDSR architecture and implemented a reference .tflite model.
- Collected and curated 25 benchmark images for OpenSR dataset.
- Delivered an invited talk:
  Mobile Super-Resolution on the MLPerf App - Benchmarking and Challenges
  Efficient Deep Learning Workshop for Computer Vision, in conjunction with **CVPR**, 2023.

Research Collaborator, Pixel-wise Adaptive Weighting for Perceptual Image Super-Resolution
Joint R&D with **Naver** (with Byeongho Heo, Ph.D., Naver AI LAB) May 2022 – May 2023

- Designed an uncertainty-based loss function for the perceptual image super-resolution model
- Achieved favorable perception-distortion tradeoff comparable to the state-of-the-art model.

Research Collaborator, Efficient Vision Transformer for Image Super-Resolution
Joint R&D with **Naver** (with Byeongho Heo, Ph.D., Naver AI LAB) May 2021 – May 2022

- Investigated and optimized vision transformer architecture for compute-efficient super-resolution.
- Reduced 75% of computation cost via a novel weighted-softmax-based attention formulation.

R&D Contributor, Generating Raw Images of Food by Domain Adaptation
Industry-Academia collaboration with **Samsung Research** Dec. 2020 – Dec. 2021

- Developed data synthesis algorithms for training vision models used in consumer electronics.
- Delivered synthetic datasets used in downstream training of consumer electronics vision models.

Internship

Student Research Intern Jan. 2019 - Jun. 2019
Research Topic: Real-World Single Image Super-Resolution
Google Cloud, Sunnyvale, CA, USA
Mentor: Ming-Hsuan Yang

Education

Ph.D. in Department of ECE Mar. 2017 – Aug. 2023
Integrated Ph.D. program in Seoul National University (SNU), Seoul, Korea
Thesis: Generalized Resampling Model for Practical Image Super-Resolution
Advisor: Kyoung Mu Lee

B.S. in Department of ECE, *Summa Cum Laude* (Rank: 9/174) Mar. 2013 – Feb. 2017
Seoul National University (SNU), Seoul, Korea

Scholarships

- **Youlchon AI Stars Scholarship**, Youlchon Foundation 2022
- **Kwanjeong Scholarship**, Kwanjeong Educational Foundation 2017 – 2018
- **National Scholarship for Science & Engineering**, Korea Student Aid Foundation 2015 – 2016
- **Scholarship of Academic Excellence**, Seoul National University 2013 – 2014

International Publications (Selected)

- Bee Lim, **Sanghyun Son**, Heewon Kim, Seungjun Nah, and Kyoung Mu Lee, "Enhanced Deep Residual Networks for Single Image Super-Resolution," **NTIRE 2017** workshop in conjunction with **CVPR**, 2017. (**Challenge winners, Workshop best paper, Over 8,000 citations on Google Scholar, over 2,500 Github stars**)

- Reyhaneh Neshatavar*, Mohsen Yavartanoo*, **Sanghyun Son**, and Kyoung Mu Lee, "ICF-SRSR: Invertible scale-Conditional Function for Self-Supervised Real-world Single Image Super-Resolution," In **WACV**, 2024.

- Joonkyu Park, **Sanghyun Son**, and Kyoung Mu Lee, "Content-Aware Local GAN for Photo-Realistic Super-Resolution," In **ICCV**, 2023.

- Wooseok Lee, **Sanghyun Son**, and Kyoung Mu Lee, "AP-BSN: Self-Supervised Denoising for Real-World Images via Asymmetric PD and Blind-Spot Network," In **CVPR**, 2022.

- Seungjun Nah, **Sanghyun Son**, Jaerin Lee, and Kyoung Mu Lee, "Clean Images are Hard to Reblur: Exploiting the Ill-Posed Inverse Task for Dynamic Scene Deblurring," In **ICLR**, 2022.

- Geonwoon Jang, Wooseok Lee, **Sanghyun Son**, and Kyoung Mu Lee, "C2N: Practical Generative Noise Modeling for Real-World Denoising," In **ICCV**, 2021.

- **Sanghyun Son** and Kyoung Mu Lee, "SRWarp: Generalized Image Super-Resolution under Arbitrary Transformation," In **CVPR**, 2021.

- **Sanghyun Son**, Jaeha Kim, Wei-Sheng Lai, Ming-Hsuan Yang, and Kyoung Mu Lee, "Toward Real-World Super-Resolution via Adaptive Downsampling Models," IEEE Trans. on Pattern Analysis and Machine Intelligence (**TPAMI**), vol. 44, no. 11, pp. 8567-8670, 2022.
https://doi.org/10.1109/TPAMI.2021.3106790

- **Sanghyun Son** and Kyoung Mu Lee, "Image Super-Resolution," in Ikeuchi K. (eds) Computer Vision. Springer, Cham, 2021. https://doi.org/10.1007/978-3-030-03243-2_838-1

- Seungjun Nah, **Sanghyun Son**, and Kyoung Mu Lee, "Recurrent Neural Networks with Intra-Frame Iterations for Video Deblurring," In **CVPR**, 2019.

- **Sanghyun Son**, Seungjun Nah, and Kyoung Mu Lee, "Clustering Convolutional Kernels to Compress Deep Neural Networks," In **ECCV**, 2018.

International Workshops and Collaborations (Selected)

- **Sanghyun Son**, Suyoung Lee, Seungjun Nah, Radu Timofte, and Kyoung Mu Lee, "NTIRE 2021 Challenge on Video Super-Resolution," **NTIRE 2021** workshop in conjunction with **CVPR**, 2021.

- **Sanghyun Son**, Jaerin Lee, Seungjun Nah, Radu Timofte, and Kyoung Mu Lee, "AIM 2020 Challenge on Video Temporal Super-Resolution," **AIM 2020** workshop in conjunction with **ICCV**, 2020.

- Seungjun Nah, **Sanghyun Son**, Radu Timofte, and Kyoung Mu Lee, "NTIRE 2020 Challenge on Image and Video Deblurring," **NTIRE 2020** workshop in conjunction with **CVPR**, 2020.

- Seungjun Nah, **Sanghyun Son**, Radu Timofte, and Kyoung Mu Lee, "AIM 2019 Challenge on Video Temporal Super-Resolution: Methods and Results," **AIM 2019** workshop in conjunction with **ICCV**, 2019.

- Seungjun Nah, Sungyong Baik, Seokil Hong, Gyeongsik Moon, **Sanghyun Son**, Radu Timofte, and Kyoung Mu Lee, "NTIRE 2019 Challenge on Video Deblurring and Super-Resolution: Dataset and Study," **NTIRE 2019** workshop in conjunction with **CVPR**, 2019.

Academic Experience & Service

Workshop Challenge Co-organizer

NTIRE 2021 Challenge on Video Super-Resolution, Video Deblurring — Jun. 2021
NTIRE 2021 workshop in conjunction with CVPR, 2021

AIM 2020 Challenge on Video Temporal Super-Resolution — Aug. 2020
AIM 2020 workshop in conjunction with ECCV, 2020

AIM 2019 Challenge on Video Temporal Super-Resolution — Sep. 2019
AIM 2019 workshop in conjunction with ICCV, 2019

Conference Reviewer

CVPR, ECCV, ICCV, and the corresponding Workshops on Image Restoration

Journal Reviewer

IEEE TPAMI, TIP, TCI
Springer IJCV
Elsevier CVIU

Awards and Honors

- **The Best Collaboration Award** from the AI Research Center, Samsung Electronics, 2024.
- **The KCCV Sang-Uk Lee Prize** (Test of Time award) from KCCV 2022.
- Winner of **Qualcomm Innovation Fellowship Korea** 2021.
- **Highly Cited Paper Award** from Department of ECE, SNU, 2018.
- **1st Place Award** in NTIRE 2017 Challenge on Single Image Super-Resolution.
- **Best Paper Award** of NTIRE 2017 Workshop: Challenge Track.

Skills

- **Programming languages**
  - **Expert:** Python
  - **Intermediate:** C++, MATLAB
  - **Novice:** CUDA (can write a custom kernel), Javascript (especially for my side projects)
- **Software stacks (selected)**
  - **General Deep Learning:** PyTorch (8+ years of experience), TensorFlow (mobile deployment)
  - **Distributed Computing:** DeepSpeed, FSDP (distributed training with >128 GPUs)
  - **Applications:** FastAPI (especially for my side projects), LangChain, vLLM, SGLang
- **Others**: Korean (Native), English (Working proficiency), Japanese (Novice), LaTeX

References

Advisor
Kyoung Mu Lee
Professor
Seoul National University
kyoungmu(at)snu.ac.kr
https://cv.snu.ac.kr/index.php/kmlee

Internship mentor
Ming-Hsuan Yang
Professor
UC Merced, Google
mhyang(at)ucmerced.edu
http://faculty.ucmerced.edu/mhyang