

# Sungmin Kang

Department of Computer Science, KAIST

[sungmin.kang@kaist.ac.kr](mailto:sungmin.kang@kaist.ac.kr) | (+82) 10-2972-7927 | GitHub ID: smkang96

Updated on December 4<sup>th</sup>, 2024

## EDUCATION

---

**Korea Advanced Institute of Science and Technology**

Mar. 2014 – Aug. 2024

- Integrated Ph.D. in Computer Science (2019 – 2024)
- B.S. in Computer Science (2014 – 2019)

## RESEARCH EXPERIENCE

---

**KAIST Computational Intelligence for Software Engineering Lab.**

Sept. 2018 – Current

Full-time Graduate Student (Mar. 2019 – Aug. 2024), Postdoctoral Researcher (Aug. 2024 – Feb. 2025)

- Thesis: Reliable Large Language Model-based Software Artifacts via Execution
- Supervisor: Dr. Shin Yoo
- Research Areas: Software Testing, Automatic Debugging

**Microsoft Research Asia**

Oct. 2022 – Apr. 2023

Research Intern

- Project: Developing an Explainable Automated Debugging Technique via Large Language Models
- Supervisor: Dr. Jian-guang Lou
- Research Areas: Language Models, Automated Program Repair

**NAVER WEBTOON Corp.**

July 2017 – Dec. 2017

Undergraduate Research Intern

- Project: Developing improved algorithms for cartoon colorization
- Supervisor: Team Leader Jaehyuk Chang, Dr. Jaegul Choo
- Research Areas: Computer Vision, Deep Learning

## PUBLICATIONS

---

### CONFERENCES

[1] Louis Milliken, **Sungmin Kang** (2<sup>nd</sup> author), Shin Yoo, “Beyond pip install: Evaluating LLM Agents for the Automated Installation of Python Projects”, *SANER* ’25, accepted as full paper.

[2] **Sungmin Kang** (co-1<sup>st</sup> author), Gabin An, Shin Yoo, “A Quantitative and Qualitative Evaluation of LLM-based Fault Localization”, *FSE* ’24, accepted as full paper.

[3] **Sungmin Kang** (co-1<sup>st</sup> author), Wonkeun Choi, Shin Yoo, “A Bayesian Framework for Automated Debugging”, *ISSTA* ’23, accepted as full paper.

[4] **Sungmin Kang** (co-1<sup>st</sup> author), Juyeon Yoon, Shin Yoo, “Large Language Models are Few-shot Testers: Exploring LLM-based General Bug Reproduction”, *ICSE* ’23, accepted as full paper.

[5] **Sungmin Kang** (1<sup>st</sup> author), Louis Milliken, Shin Yoo, “Detecting Inaccurate Descriptions in LLM-generated Code Comments”, preprint.

## JOURNALS

- [1] **Sungmin Kang** (1<sup>st</sup> author), Bei Chen, Shin Yoo, Jian-guang Lou, “Explainable Automated Debugging via Large Language Model-driven Scientific Debugging”, *Empirical Software Engineering*, accepted.
- [2] **Sungmin Kang** (co-1<sup>st</sup> author), Juyeon Yoon, Nargiz Askarbekkyzy, Shin Yoo, “Evaluating Diverse Large Language Models for Automatic and General Bug Reproduction”, *ACM Transactions on Software Engineering*, accepted.
- [3] **Sungmin Kang** (1<sup>st</sup> author), Robert Feldt, Shin Yoo, “Deceiving Humans and Machines Alike: Search-based Test Input Generation for DNNs using Variational Autoencoders”, *ACM Transactions on Software Engineering and Methodology*, accepted.
- [4] Jeongju Sohn, **Sungmin Kang** (co-1<sup>st</sup> author), Shin Yoo, “Arachne: Search Based Repair of Deep Neural Networks”, *ACM Transactions on Software Engineering and Methodology*, accepted.
- [5] Kyeong Min Song, Shinho Kim, **Sungmin Kang** (3<sup>rd</sup> author), Tae Won Nam, Geon Yeong Kim, Hunhee Lim, Eugene N. Cho, Kwang Ho Kim, Se Hun Kwon, Min Seok Jang et al., “Microcellular sensing media with ternary transparency states for fast and intuitive identification of unknown liquids”, *Science Advances*, accepted.
- [6] Koeun Han, Hee-Jin Jeong, Hee-Bum Yang, **Sung-Min Kang** (4<sup>th</sup> author), Jin-Kyung Kwon, Seungill Kim, Doil Choi, and Byoung-Cheorl Kang, “An ultra-high-density bin map facilitates high-throughput QTL mapping of horticultural traits in pepper”, *DNA Research*, accepted.

## WORKSHOPS AND SHORT PAPERS

- [1] Jae Yong Lee, **Sungmin Kang** (2<sup>nd</sup> author), Juyeon Yoon, Shin Yoo, “The GitHub Recent Bugs Dataset for Evaluating LLM-based Debugging Applications”, *ICST’24*, accepted to Demonstration track.
- [2] Robert Feldt, **Sungmin Kang** (2<sup>nd</sup> author), Juyeon Yoon, Shin Yoo, “SOCRATEST - Towards Autonomous Testing Agents via Conversational Large Language Models”, *ASE’23*, accepted to NIER track.
- [3] **Sungmin Kang** (1<sup>st</sup> author), Shin Yoo, “GLAD: Neural Predicate Synthesis to Repair Omission Faults”, *ICSE’23*, accepted as poster.
- [4] **Sungmin Kang** (1<sup>st</sup> author), Shin Yoo, “Towards Objective-Tailored Genetic Improvement Through Large Language Models”, *GI’23*, accepted as position paper.
- [5] **Sungmin Kang** (1<sup>st</sup> author), Shin Yoo, “Language Models Can Prioritize Patches for Practical Program Patching”, *ICSE’22 Workshop on Automated Program Repair*, 2022, accepted as full paper.
- [6] **Sungmin Kang** (1<sup>st</sup> author), Shin Yoo, “Improving Fault Localization and Automated Program Repair with Suspicious Predicates”, *KCSE’22*, accepted as short paper.
- [7] **Sungmin Kang** (1<sup>st</sup> author), Robert Feldt, Shin Yoo, “SINVAD: Search-based Image Space Navigation for DNN Image Classifier Test Input Generation”, *ICSE’20 Workshop on Search-based Software Testing*, 2020, accepted as full paper.
- [8] **Sungmin Kang** (1<sup>st</sup> author), Jaegul Choo, and Jaehyuk Chang, “Consistent Comic Colorization with Pixel-wise Background Classification”, *NIPS’17 Workshop on Machine Learning for Creativity and Design*, 2017, accepted as poster.

## PATENTS

---

- Method for coloring a target image, and device and computer program therefor, US Patent App. 17/464,899
- Method, apparatus, and computer program for completing painting of image, and method, apparatus, and computer program for training artificial neural network, US Patent 11887224
- Method and apparatus for operating benchmark system of software-debugging, Korean Patent App. 10-2024-0087100

## INVITED TALKS

---

- “Explainable Program Repair with Automated Scientific Debugging”.  
Talk expected at Dagstuhl Seminar on Automated Programming and Program Repair (2024).
- “Towards Reliable LLM-based Software Generation and Debugging”.  
Talk at NUS (2024) and Georgia Tech (2024).
- “The State-of-the-art in LLM-based SE techniques: From Bug Reproduction to Program Repair”.  
Talk given at SAP Korea and SureSoft (2023).

## SERVICE

---

### JOURNAL REVIEWER

Reviewer for TSE (2023, 2024), TOSEM (2023, 2024), JSS (2024), ASE (2024), IEEE Software (2023), PLOS ONE (2024), IEEE Communications (2024)

### CONFERENCE/WORKSHOP REVIEWER

PC member for ICSE (2026), ISSRE (2024), GI (2024, 2025), SBFT (2025), DeepTest (2025), LLMfwEC (2024)

## AWARDS

---

- |  |      |
|--|------|
| <b>Best Position Paper Award (International Workshop on Genetic Improvement 2023)</b>                | 2023 |
| For the paper “Towards Objective-Tailored Genetic Improvement Through Large Language Models”         |      |
| <b>Best Presentation Award (International Workshop on Genetic Improvement 2023)</b>                  | 2023 |
| For the paper “Towards Objective-Tailored Genetic Improvement Through Large Language Models”         |      |
| <b>Best Short Paper Award (Korea Conference on Software Engineering 2022)</b>                        | 2022 |
| For the paper “Improving Fault Localization and Automated Program Repair with Suspicious Predicates” |      |
| <b>Excellent Teaching Assistant Reward (KAIST School of Computing)</b>                               | 2019 |
| As an assistant of the “AI Based Software Engineering” course  |      |

## SKILLS

---

### Natural Languages

Korean (native), English (native)

### Programming Languages

Fluent in Python, familiar with C, bash.