Seongjae Kang

Homepage | LinkedIn | GitHub | Google Scholar | Twitter

Seoul, South Korea

RESEARCH INTERESTS

I am passionate about developing innovative AI solutions that address real-world challenges across diverse domains, with a commitment to making a meaningful impact on society by bridging the gap between AI and people through both research and engineering. My current interests involve:

- **Training task-specific AI models**: Leveraging the growing capabilities of foundation models in data-efficient settings, utilizing techniques such as knowledge distillation (KD), few-shot learning, self-supervised and semi-supervised learning, and active learning (AL).
- **Building automated systems**: Developing systems that genuinely automate and assist human workflows by integrating AI models such as large language models (LLMs), multimodal large language models (MLLMs) through instruction tuning, reinforcement learning (RL), and agentic AI frameworks.

EXPERIENCE

VUNO **May 2024 – Present**

Research Scientist/Engineer (Military Alternative Service) *Seoul, South Korea*

- **Medical Image Analysis:** Developed deep learning models for detecting findings and diseases in chest X-ray images, investigating methods to leverage chest X-ray foundation models' knowledge for compact deployment-ready models
  - Developed methods for data-efficient learning approaches including few-shot learning and semi-supervised learning, tailored to the characteristics of specific findings and diseases
  - Conducted research on knowledge distillation framework to transfer large vision-language medical foundation models into compact models suitable for clinical deployment
- **Human-in-the-Loop Data Acquisition:** Designed and implemented active learning pipeline to address the high labeling costs of medical imaging data
  - Developed full-stack proof-of-concept data engine integrating Label Studio and MongoDB for efficient annotation workflows
  - Conducted research on active learning frameworks and selection method under the regime of foundation model knowledge distillation
- **Automated Radiology Report Generation:** Developed AI systems to assist radiologist workflows through automated report generation
  - Investigated how recent Multimodal Large Language Models (MLLMs) incorporate clinical contexts (e.g., indication, clinical history) when generating radiology reports
  - Developed proof-of-concept demonstration for keyword-based radiology report generation to assist radiologist workflows using MLLMs and Gradio UI
- **Workflow Automation:** Voluntarily developed automated cash reimbursement system that automatically fills submission forms, reducing processing time by approximately 1 hour per month for each employee
  - Implemented JavaScript-based web automation for form interaction and navigation
  - Designed user interface and distributed solution via Chrome Extension Store

Deepauto AI **Feb 2024 – Mar 2024**
Research Scientist/Engineer *Seoul, South Korea*

- Conducted research on enabling Multimodal Large Language Models (MLLMs) to tackle a variety of computer vision tasks as universal task solver
  - Reformulated image classification, object detection, image segmentation into instruction tuning dataset
  - Formulated image tokens from VQVAE to be integrated into LLaVA framework for image output generation

VoyagerX **Sep 2020 – Sep 2021**
Research Engineer (Internship) *Seoul, South Korea*

- Developed deep learning models for document quality enhancement and finger removal from scanned documents
  - Collected diverse datasets including papers with various lighting conditions and materials, and finger images from various postures and demographics
  - Created synthetic datasets to train supervised models for both document enhancement and finger removal tasks
  - Deployed lightweight models with TensorFlow Lite for efficient edge device inference and integrated into Android app using Android Studio and Java

SGVR Lab, KAIST **Mar 2019 – Aug 2019**
Undergraduate Research Assistant *Daejeon, South Korea*
Advisor: Prof. Sung-eui Yoon

- Developed virtual piano system that generates musical output by analyzing performer's hand movements through computer vision
  - Collected video dataset capturing finger movements on piano keys from multiple participants
  - Designed spatial-temporal model to detect key locations and pressing intensity from visual information
  - Integrated MIDI synthesis to generate realistic piano sounds based on detected movements

Computer Vision Lab, KAIST **Jun 2017 – Jun 2018**
Undergraduate Research Assistant *Daejeon, South Korea*
Advisor: Prof. Hyunseung Yang

- Conducted research on computer vision and deep learning methodologies

EDUCATION

KAIST (Korea Advanced Institute of Science and Technology) **Feb 2022 – Feb 2024**
M.S. in Graduate School of AI *Seoul, South Korea*
Advisor: Prof. Sung Ju Hwang
GPA: 3.96/4.3

KAIST (Korea Advanced Institute of Science and Technology) **Feb 2016 – Feb 2022**
B.S. in Computer Science (major) and Electrical Engineering (minor) *Daejeon, South Korea*
GPA: 3.88/4.3

AWARDS & HONORS

Korea National Science & Technology Scholarship **2018 – 2021**
Korea Ministry of Science and ICT *South Korea*

1st Prize, Semantic Segmentation AI Model for Body Parts Challenge **2021**
Alchera, AI Heroes *South Korea*

Linux Master 1st Grade Certificate
Korea Information Technology Human Resources Development Institute *South Korea*

PUBLICATIONS

*(P: preprint, C: conference, J: journal, W: workshop, *: equal contribution)*

- **[P1] Automated Structured Radiology Report Generation with Rich Clinical Context**
  Seongjae Kang*, Dong Bok Lee*, Juho Jung, Dongseop Kim, Won Hwa Kim, Sunghoon Joo
  under review, 2025

- **[P2] PCoreSet: Effective Active Learning through Knowledge Distillation from Vision-Language Models**
  Seongjae Kang*, Dong Bok Lee*, Hyungjoon Jang, Dongseop Kim, Sung Ju Hwang
  under review, 2025

- **[P3] Simple yet Effective Semi-supervised Knowledge Distillation from Vision-Language Models via Dual-Head Optimization**
  Seongjae Kang*, Dong Bok Lee*, Hyungjoon Jang, Sung Ju Hwang
  under review, 2025

- **Virtual Piano using Computer Vision**
  Seongjae Kang, Jaeyoon Kim, Sung-eui Yoon
  arXiv preprint arXiv:1910.12539, 2019

PERSONAL PROJECTS

YouTube Multi-Subtitle for Chinese Learning
Tools: JavaScript, Chrome Extension API

- Built a web application for learning Chinese through YouTube videos with simultaneous display of three subtitle tracks
  - Recognized that Chinese requires both characters and Pinyin (pronunciation) to learn effectively, unlike other languages where a single subtitle suffices
  - Implemented real-time Chinese-to-Pinyin conversion to automatically generate pronunciation guides
  - Enabled loading of English and Chinese subtitles from multiple sources including machine translation
  - Replaced YouTube's default subtitle with custom overlay showing English meaning, Chinese characters, and Pinyin together
- Made Chinese learning significantly easier by providing all necessary linguistic information in one place

SKILLS

Languages: Korean (Native), English (Fluent), Spanish (● ● ○), Chinese (● ○ ○)
Programming Languages: Python, Java, C, JavaScript, and English*
Interests: Continuous Self-improvement, Language Learning, and Swimming

Last updated on November 12, 2025