MINSEOK KANG

Computer Vision Engineer, ML/DL Researcher

+82 10-4228-0491 louis0503@yonsei.ac.kr github.com/minseokii minseokii.github.io

RESEARCH INTERESTS

Scene Understanding: Video Temporal Grounding, Video Scene Graph Generation

Vision–Language Multimodal Learning: Video Large Language Model

EDUCATION

Integrated M.S./Ph.D. Mar. 2022 – Feb. 2027 (expected)
Yonsei University, Image and Video Pattern Recognition Lab. Seoul, Korea
Electrical and Electronic Engineering, GPA 4.07/4.3

Bachelor of Science Mar. 2017 – Feb. 2021
Korea University Seoul, Korea
Electrical and Electronic Engineering, GPA 3.71/4.5 (Major 4.13/4.5)

Kyunggi High School Mar. 2014 – Feb. 2017
Seoul, Korea

PUBLICATIONS

OTT-Vid: Optimal Transport Temporal Token Compression for Video Large Language Models
Minseok Kang, Minhyeok Lee, Jungho Lee, Minjung Kim, Donghyeong Kim, Dayeon Lee, Heeseung Choi, Ig-Jae Kim, Sangyoun Lee — *Under Review at NeurIPS 2026*

Revisiting Weakly-Supervised Video Scene Graph Generation via Pair Affinity Learning
Minseok Kang, Minhyeok Lee, Minjung Kim, Jungho Lee, Donghyeong Kim, Sungmin Woo, Inseok Jeon, Sangyoun Lee — *ECCV 2026*

SwiftVGGT: A Scalable Visual Geometry Grounded Transformer for Large-Scale Scenes
Jungho Lee, Minhyeok Lee, Sunghun Yang, **Minseok Kang**, Sangyoun Lee — *CVPR Findings 2026*

Seen-to-Scene: Keep the Seen, Generate the Unseen for Video Outpainting
Inseok Jeon, Minhyeok Lee, Seunghoon Lee, **Minseok Kang**, Suhwan Cho, Sangyoun Lee — *CVPR Findings 2026*

Generalizing CLIP Prompts for Zero-Shot Anomaly Detection
Donghyeong Kim, Chaewon Park, Suhwan Cho, Hyeonjeong Lim, **Minseok Kang**, Jungho Lee, Sangyoun Lee — *Pattern Recognition 2026*

Empower Words: DualGround for Structured Phrase and Sentence-Level Temporal Grounding
Minseok Kang, Minhyeok Lee, Minjung Kim, Donghyeong Kim, Sangyoun Lee — *NeurIPS 2025*

CMTM: Cross-Modal Token Modulation for Unsupervised Video Object Segmentation
Inseok Jeon, Suhwan Cho, Minhyeok Lee, Seunghoon Lee, **Minseok Kang**, Jungho Lee, Chaewon Park, Donghyeong Kim, Sangyoun Lee — *ICIP 2025*

Multilevel Artificial Electronic Synaptic Device of Direct Grown Robust MoS_2 Based Memristor Array for In-Memory Deep Neural Network
Muhammad Naqi*, **Minseok Kang***, Na Liu*, Taehwan Kim, Seungho Baek, Arindam Bala, Changgyun Moon, Jongsun Park, Sunkook Kim — *npj 2D Materials and Applications, 6, 53, Aug. 2022*

* indicates equal contribution.

PROJECTS

Robust Detection of Abnormal and Hazardous Behaviors under Nighttime and Adverse Weather Conditions Jul. 2024 – Feb. 2027

Yonsei University | Funded by Electronics and Telecommunications Research Institute (ETRI)

- Project Leader / Deep Learning Researcher
- Construction of multi-sensor dataset integrating RGB, Depth, Thermal, IR, and LiDAR modalities
- Multi-sensor fusion for robust object and behavior recognition under low-visibility conditions
- Development of an efficient intelligent image processing algorithm resilient to complex degradations

Collaborative Perception and Intelligence Framework for Hyper-connected Interaction among Human and Intelligent Things Apr. 2024 – Dec. 2025

Korea Electronics Technology Institute (KETI)

- Deep Learning Researcher
- Development of an efficient skeleton-based action recognition model

Domain Generalization-Based Face Recognition Anti-Spoofing Model Aug. 2023 – Aug. 2024

Yonsei University | Funded by Samsung Electronics, Mobile eXperience (MX)

- Deep Learning Researcher
- Construction of a multi-device facial image database captured under diverse illumination and pose conditions
- Development of a domain generalized face anti-spoofing model robust across diverse camera domains

2D-3D Feature Correspondence Learning for Virtual Scene Reconstruction Jun. 2022 – Jun. 2023

Yonsei University | Funded by LG Electronics, Business Solutions (BS)

- Deep Learning Researcher
- Development of core technology for constructing a virtual 3D space by estimating 3D position and pose
- Development of stable 3D scene reconstruction from monocular 2D image

PATENTS

Method and Apparatus for Multi-domain Object Detection Based on Geometric Information 2026

KR Patent 10-2026-0026131

Multi-Domain Based Device and Method for Object Detection and Tracking 2025

KR Patent 10-2025-0024051

Electronic Device and Method for Detecting Forgery of Biometric Information 2024

KR Patent 10-2024-0123340

HONORS AND AWARDS

3rd Place, Perception Test Challenge 2025 — Task 3 2025

Temporal Action Localization & Temporal Sound Localization, Hosted by DeepMind (ICCV Workshop 2025)

Semester High Honors, School of Engineering, Korea University Spring 2019, Fall 2019, Spring 2020

Recognition for top academic performance

Top-Honor Scholarship, School of Electrical Engineering, Korea University 2019, 2020

Merit-based scholarship for outstanding academic achievement

SKILLS

Languages: Korean (Native), English (Professional working proficiency)

Programming (Advanced): Python, PyTorch, Linux OS, Bash/Shell scripting

Programming (Intermediate): C/C++, MATLAB, Verilog