

MINJUNG KIM

✉ minjung.kim@vision.snu.ac.kr | 🎓 Google scholar | 💬 LinkedIn | 🐣 GitHub | 🏠 minnjung.github.io
📍 317, Bldg 138, 1 Gwanak-ro, Gwanak-gu, Seoul, Korea

RESEARCH INTERESTS

My research focuses on **Visual Localization** and **3D Dense Captioning** for enhanced 3D scene understanding, with particular interest in: (i) understanding complex scenes from images and point clouds, (ii) effectively handling multi-modalities, and (iii) achieving a comprehensive understanding of 3D scenes through natural language.

EDUCATION

Seoul National University

Integrated M.S./Ph.D. Student in Computer Science and Engineering; (GPA: 4.03/4.3)
Vision and Learning lab, advised by Prof. Gunhee Kim; **Outstanding Doctoral Thesis Award**

Seoul, Korea

Mar. 2018 – Feb. 2025

Sogang University

B.S. in Computer Science and Engineering; (GPA: 3.58/4.3), **Magna Cum Laude**
Advised by Prof. Hyukjun Lee

Seoul, Korea

Mar. 2014 – Feb. 2018

PUBLICATIONS

Bi-directional Contextual Attention for 3D Dense Captioning

ECCV 2024

Minjung Kim, Hyung Suk Lim, Soonyoung Lee, Bumsoo Kim*, Gunhee Kim*

Oral presentation

Rethinking the Role of Queries in 3D Dense Captioning

KCC 2024

Minjung Kim, Gunhee Kim

See It All: Contextualized Late Aggregation for 3D Dense Captioning

ACL 2024 Findings

Minjung Kim, Hyung Suk Lim, Seung Hwan Kim, Soonyoung Lee, Bumsoo Kim*, Gunhee Kim*

EP2P-Loc: End-to-End 3D Point to 2D Pixel Localization for Large-Scale Visual Localization

ICCV 2023

Minjung Kim, Junseo Koo, Gunhee Kim

Indoor/Outdoor Transition Recognition Based on Door Detection

UR 2022

Seohyun Jeon, Minjung Kim, Seunghwan Park, Jaeyoung Lee

Drop-Bottleneck: Learning Discrete Compressed Representation for Noise-Robust Exploration

ICLR 2021

Jaekyeom Kim, Minjung Kim, Dongyeon Woo, Gunhee Kim

Logo Detection and Recognition Algorithm using YOLO-v3 Model

CICS 2020

Minjung Kim, Sungen Kim, Gunhee Kim

Memorization Precedes Generation: Learning Unsupervised GANs with Memory Networks

ICLR 2018

Youngjin Kim, Minjung Kim, Gunhee Kim

Machine Learning for Determining Duplicate Question

KSC 2017

Minjung Kim, Yeongjoon Park, Hyung Suk Lim, Jihoon Yang

Sketch based Face Image Generation with Text Mapping

KSC 2017

Minjung Kim, Hyung Suk Lim, Yeongjoon Park, Yeseul Joo, Myoung Wan Koo

EXPERIENCES

| | |
|--------------------------------------------------------------|------------------------------|
| Vision Lab | LG AI Research |
| <i>AI Scientist</i> | <i>Jun. 2025 – Current</i> |
| Vision and Learning Lab | Seoul National University |
| <i>Postdoctoral Researcher</i> | <i>Feb. 2025 – Jun. 2025</i> |
| Vision and Multimodal Lab | LG AI Research |
| <i>Research Intern</i> | <i>Jun. 2023 – May. 2024</i> |
| KDB-SNU AI course | Seoul National University |
| <i>Teaching Assistant</i> | <i>Apr. 2023</i> |
| 2022-3 SK hynix ML Engineer course | Seoul National University |
| <i>Teaching Assistant</i> | <i>Nov. 2022 – Dec. 2022</i> |
| KDB-SNU AI course | Seoul National University |
| <i>Teaching Assistant</i> | <i>Apr. 2022 – May. 2022</i> |
| LG AI core human resource training course | Seoul National University |
| <i>Teaching Assistant</i> | <i>Feb. 2022</i> |
| IoT · Artificial Intelligence · Big Data (IAB) course | Seoul National University |
| <i>Teaching Assistant</i> | <i>Sep. 2018 – Jun. 2019</i> |
| Bayesian Deep Learning course | Boostcourse, Naver Connect |
| <i>Publisher</i> | <i>Feb. 2018 – Jul. 2018</i> |
| Vision and Learning Lab | Seoul National University |
| <i>Research Intern</i> | <i>Jul. 2017 – Feb. 2018</i> |
| Biointelligence Laboratory | Seoul National University |
| <i>Research Intern</i> | <i>Sep. 2016 – Feb. 2017</i> |
| Arduino & Raspberry Pi Kit Developer | MakeWith (Startup) |
| <i>Development Intern</i> | <i>Dec. 2016 – Jan. 2017</i> |

AWARDS & SCHOLARSHIPS

| | |
|----------------------------------------------------------------------------------------------------------------|----------------------------------------------------------------------|
| Outstanding Doctoral Thesis Award | Dept. of Computer Science and Engineering, Seoul National University |
| <i>Academic Honors</i> | <i>Feb. 2025</i> |
| Youlchon AI Star | Youlchon Foundation, Nongshim Group |
| <i>Fellowship</i> | <i>Sep. 2024</i> |
| Animal Datathon Korea | Animal Tech Korea |
| Predicting joint coordinates of a cow for pose estimation; 2nd place | <i>Jul. 2021</i> |
| The 27th Samsung Humantech Paper Award | Samsung Electronics |
| Signal Processing section; Silver prize | <i>Feb. 2021</i> |
| Magna Cum Laude Honor | Sogang University |
| <i>Academic Honors</i> | <i>Feb. 2018</i> |
| KSC 2017 Paper Award | Korean Institute of Information Scientists and Engineers |
| The Undergraduate/Junior Thesis Contest Award | <i>Feb. 2018</i> |
| Academic Excellence Scholarship | Sogang University |
| <i>Academic Honors</i> | <i>Jul. 2017 – Feb. 2018</i> |
| Windows 10 IoT Core & Microsoft Azure for Microsoft IoT Solution Competition | Microsoft |
| Implementing Internet of Things (IoT) projects with Windows 10 IoT Core and Microsoft Azure; 10th place | <i>Apr. 2017</i> |

PROJECTS

DeepGuider | [GitHub](#)

Apr. 2019 – May. 2023

- The DeepGuider Project is a national government-funded research project focused on developing a navigation guidance system that enables robots to navigate urban environments without the need for pre-mapping.
- I contributed by identifying clues to help locate autonomous robots, detecting and recognizing points of interest (POIs) in scene images, including text, landmarks, and doors for indoor-outdoor transitions, while also developing robust training methods to adapt to environmental changes.

PRIDE: 3D Place Recognition In Dynamic Environment | [GitHub](#)

Mar. 2022 – Apr. 2023

- This work proposes a new dataset called PRIDE, which includes dynamic objects such as cars and pedestrians, for 3D place recognition in dynamic environments that are more realistic and challenging than current benchmarks.
- The proposed PRIDE-Net architecture with a new loss function focuses on extracting discriminative global descriptors and capturing global context using spatial information, while being robust to dynamic environments.
- Experiments on the PRIDE dataset and existing benchmarks show that our proposed method outperforms previous methods and that each proposed module effectively improves performance.
- The code will be released after acceptance.

FCAT: Fully Convolutional Network with Self-Attention for Point Cloud based Place Recognition | [GitHub](#)

Dec. 2020 – Feb. 2022

- We construct a novel network named FCAT (Fully Convolutional network with a self-ATTention unit) that can generate a discriminative and context-aware global descriptor for place recognition from the 3D point cloud.
- It features with a novel sparse fully convolutional network architecture with sparse tensors for extracting informative local geometric features computed in a single pass. It also involves a self-attention module for 3D point cloud to encode local context information between local descriptors.

Bayesian Deep Learning course | [Lecture](#)

Feb. 2018 – Jul. 2018

- To understand deep learning papers, we explain the basic concepts of probability and Bayesian, and introduce papers related to Bayesian neural networks.
- This lecture can be taken through *edwith* of Naver Connect.

Sketch based Face Image Generation with Text Mapping | [GitHub](#)

Sep. 2017 – Feb. 2018

- A typical sketch might have been uncomfortable when a person or program was used to map a person's features in detail. This process is limited not only because it is very complex and requires technicians, but also because it creates a feeling of incompatibility with real people.
- This program, named Metamon, makes a picture of a person's face by entering the image of the border sketch of the person's face and the text information that shows the characteristics of the face.

Arduino & Raspberry Pi & Internet of Things (IoT) Tutorial | [Project page](#)

Dec. 2016 – Mar. 2017

- I create tutorial pages with Youtube videos and code for beginners in Arduino kit and Raspberry Pi development.
- I also introduce the concept of the Internet of Things (IoT) and work on a mini-project using *ThingSpeak*™.

Sogang Navigation and Introduction (SNI) | [Github](#)

Mar. 2015 – Jul. 2015

- We develop a navigation system that introduces the internal facilities of each building and displays the shortest route and time from building to building using the Floyd-Washall algorithm.
- To build data for the development, we measured the time taken by walking directly on each path.

SKILLS

Programming: Python, C, C++

Frameworks: Pytorch, TensorFlow/Keras

Tools: Git, VSCode, Vim, Docker, Slurm

Others: Arduino, Rapsberry Pi

PROFESSIONAL ACTIVITES

Reviewer of International Conferences

- European Conference on Computer Vision (ECCV) 2024
- IEEE/CVF International Conference on Computer Vision (ICCV) 2023, 2025
- IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR) 2023, 2025
- Asian Conference on Computer Vision (ACCV) 2022
- International Conference on Learning Representations (ICLR) 2022, 2023
- Neural Information Processing Systems (NeurIPS) 2021, 2022, 2023, 2024

Reviewer of International Journals

- International Journal of Computer Vision (IJCV) 2024

Technical Coaching

- 2022-3 SK hynix ML Engineer Technical Coaching