Junho Park

✉ junho18.park@gmail.com
in junho18-park
🌐 https://redorangeyellowy.github.io/

Introduction

I am an AI researcher specializing in **generative models, vision foundation models, physical AI, and on-device AI**. I have led large-scale industrial and academic projects, consistently achieving world-class performance and publishing my work in top-tier conferences and journals such as **ECCV (Oral), ICCV, IEEE TMM, IEEE TIM, and IEEE Access**. Through interdisciplinary collaboration with global industry leaders **(LG, Samsung)** and academic institutions **(Oxford, ETH Zurich)**, I have developed a broad perspective and a proven ability to drive impactful research.

Employment

2024 – · · · ·	**AI Researcher,** AI Lab, LG Electronics. (*Leader: Ph. D. Jaechul Kim*)

Education

2022 – 2024	**M.S., Sogang University,** Electronic Engineering. (*Advisor: Prof. Suk-Ju Kang*)
2014 – 2022	**B.S., Sogang University,** Mathematics and Electronic Engineering (Double Major).

Technical Stacks

Generative AI	Large Language Models (LLM), Diffusion Models.
Vision AI	Vision Foundation Models (VFM), Object Detection, Panoptic Segmentation, Pose/Gaze Estimation, Depth Estimation, Face Recognition, Person Re-Identification.
Physical AI	Egocentric Vision, Hand-Object Interaction.
On-Device AI	Lightweight Modeling, Quantization.
Industrial AI	Image Restoration, Few-Shot Learning, Machine Learning.

Work Experiences

Large-Scale Generative Datasets, 2025.3–

Affiliation	**LG Electronics.**
Goal	**42M-Scale Large Generative Datasets for Vision Tasks.**
Responsibilities	Building a generation pipeline based on a state-of-the-art diffusion model to maximize the reliability of generated datasets.
	Aiming to reduce the performance gap to within approximately 2% compared to training on real datasets.

Vision Foundation Models, 2024.3–2025.2

Affiliation	**LG Electronics.**
Goal	**World-Best Performance for General Vision Tasks.**

Work Experiences (continued)

Responsibilities
- Developed a multi-task transformer-based Vision Foundation Model (VFM) that performs object detection, segmentation, pose, and depth estimation simultaneously.
- Developed an on-device dedicated lightweight VFM based on CNNs for low-cost chips.
- Developed task-specific pseudo-labellers (e.g., Grounding-DINO, SAM, PCT, Depth Anything) to build a large-scale database.

Performance
- Achieved world-best performance (Detection: AP 5.19%↑, Segmentation: AP 9.41%↑, Pose Estimation: AP 3.09%↑, Depth Estimation: Abs-Rel 1.88%↓)
- Achieved 40 FPS on the target chip (Snapdragon Gen2).
- Reduced the number of parameters in the lightweight VFM by approximately 50% and achieved 200 FPS on the target chip.
- Showcased as the representative Vision AI of LG at CES 2025.

Research Experiences

LLM-based 3D Room Generation, 2023.7–2025.2

Affiliation
- **Sogang University w/ Pusan National University.**

Responsibilities
- Claimed the limitations in existing 3D scene generation models (e.g., Text2Room, Holodeck, SceneScape), which fail to capture the complexity of indoor environments.
- Developed a model that sequentially constructs various room elements (e.g., texture, furniture, shape, etc.) and allows user-friendly, interactive editing.
- Built a model that modularizes room components like a Python program, based on LLM-driven text-to-code generation.
- Developed a diffusion-based panoramic image generation model specialized for generating room shapes and textures.

Performance
- Improved panoramic image generation performance over state-of-the-art methods (FID: 29.3%↓, KID: 61.9%↓).
- Achieved better 3D room generation performance compared to state-of-the-art approaches (PQ: 33.2%↑, 3DS: 59.8%↑).

Publication
- **IEEE TMM 2025, Programmable-Room: Interactive Textured 3D Room Meshes Generation Empowered by Large Language Models. (link)**
- **ECCVW 2024 Oral, Interactive 3D Room Generation for Virtual Reality via Compositional Programming. (link)**

Diffusion-based 3D Hand Datasets Generation, 2023.7–2024.6

Affiliation
- **Sogang University w/ Pusan National University.**

Responsibilities
- Developed a hand image generation model to complement Stable Diffusion, which generally struggles with generating accurate hand images.
- Proposed a global & local visual prompt-based method inspired by the observation that hands typically occupy only a small region in images.
- Introduced a technique to enhance hand visibility by focusing on hand-related tokens (e.g., holding, grasping, or hand) in the input text.
- Extended the approach to create a 3D hand dataset aimed at improving the performance of downstream tasks (e.g., 3D hand reconstruction).

Research Experiences (continued)

Performance	◪ Improved hand image generation performance over state-of-the-art methods (FID: 4.63%↓, KID: 9.97%↓, MSE: 15.28%↓).
	◪ Enhanced 3D hand reconstruction performance (MPVPE: 3.80%↓, RRVE: 1.17%↓, MR-RPE: 6.40%↓).
Publication	◪ **ECCV 2024 Oral, AttentionHand: Text-driven Controllable Hand Image Generation for 3D Hand Reconstruction in the Wild. (link)**
	◪ **ICCVW 2023, A Novel Framework for Generating In-the-Wild 3D Hand Datasets. (link)**

Robot Learning w/ Exocentric-to-Egocentric View Translation, 2024.10–

Affiliation	◪ **Independent Researcher w/ VGG (Oxford) and ETH Zurich.**
Responsibilities	◪ Demonstrated exocentric-to-egocentric view translation remains challenging due to significant visual and geometric differences between perspectives.
	◪ Proposed a novel framework leveraging rich exocentric observations to translate views.
	◪ Developed a two-stage pipeline extracting diverse cues, including projected point clouds, 3D hand poses, and textual descriptions, and inpainting egocentric views from them.
Performance	◪ Achieved state-of-the-art performance on the H2O and TACO benchmarks (FID: 44.23%↓, PSNR: 16.84%↑, SSIM: 10.53%↑, LPIPS: 23.42%↓).
	◪ Demonstrated strong generalization in unseen scenarios involving unseen objects, actions, scenes, and subjects.
Publication	◪ **Under Review, EgoWorld: Translating Exocentric View to Egocentric View using Rich Exocentric Observations.**

Robot Learning w/ Hand-Object Image Translation, 2024.3–

Affiliation	◪ **Sogang University.**
Responsibilities	◪ Addressed the challenges of the image translation for hand-object interaction, which requires understanding spatial relationships and inferring pixel-level appearance from new viewpoints.
	◪ Built a new state-of-the-art model to generate target images from a source image and target description, using a pose estimator, pose translator, pose projector, and scene translator.
	◪ Designed the scene translator with state-of-the-art diffusion models to fully leverage appearance and pose information.
Performance	◪ Achieved state-of-the-art performance across various hand-object interaction scenarios, including H2O, DexYCB, and HO3D (FID: 29.59%↓, KID: 57.64%↓, LPIPS: 24.54%↓, SSIM: 1.832%↑, PSNR: 4.482%↑).
	◪ Demonstrated that our model can be applied to various downstream tasks across multiple applications, including seamless video generation and hand-hand interaction.
Publication	◪ **Under Review, Describe Your Camera: Towards Implicit 3D-Aware Image Translation for Hand-Object Interaction.**
	◪ **ECCVW 2024, Diffusion-based Interacting Hand Pose Transfer. (link)**

Few-Shot SEM Image Restoration & Machine Learning, 2023.3–2024.2

Affiliation	◪ **Sogang University w/ Samsung Electronics.**

Research Experiences (continued)

Responsibilities

- Applied the image denoising techniques to remove noise from scanning electron microscope (SEM) images acquired under various conditions such as different energy levels and currents.
- Applied the super-resolution techniques to upsample low-resolution SEM images to high-resolution.
- Developed a regression model to predict structural lengths from the restored SEM images.
- Developed an few-shot data augmentation algorithm to compensate for the limited training data inherent to the domain.

Performance

- Successfully deployed SEM image denoising and super-resolution under environments with various conditions.
- Achieved improved performance in structural length prediction compared to state-of-the-art architectures such as ViT and ConvNeXt(v2) (MAE: 80.5% ↓, R2: 10.1% ↓, MAPE: 80.7% ↓)

Publication

- **IEEE TIM 2024, Mixup-based Neural Network for Image Restoration and Structure Prediction from SEM Images. (link)**

Calibration-Free Gaze Tracking in Public Space, 2022.3–2023.2

Affiliation

- **Sogang University w/ Korea Electronics Technology Institute (KETI).**

Responsibilities

- Developed a gaze tracking algorithm that operates on large displays over 55 inches.
- Developed a calibration-free algorithm suitable for public spaces, enabling use by general users rather than specific individuals.
- Built an on-device model capable of running in real-time for autonomous driving environments.

Performance

- Achieved performance improvement over existing algorithms (F1-score: 12.4%↑).
- Achieved 30 FPS on the target chip (NVIDIA Jetson Xavier).
- Deployed the model on autonomous vehicles such as Navya and Robo.

Publication

- **IEEE Access 2023, Improving Gaze Tracking in Large Screens with Symmetric Gaze Angle Amplification and Optimization Technique. (link)**

Skills

Languages

- Strong reading, writing and speaking competencies for English and Korean.

Machine Learning Libraries

- PyTorch, Huggingface, PyTorch Lightning, Tensorflow, …

Data Analytics

- Numpy, Matplotlib, SciPy, Pandas, Seaborn, …

General Software Engineering

- Python, Git-based workflow, CUDA, Shell, Linux, …

Miscellaneous

Awards

2021

- **1st place in AI Grand Challenge.** (link)

Miscellaneous (continued)

Academic Service

2025–   ICCV Reviewer.

             IEEE TCSVT Reviewer.

2024–   IEEE TII Reviewer.