Woo Jae Kim

AI robustness · AI safety · Adversarial attacks & defense

School of Computing
Korea Advanced Institute of Science and Technology (KAIST)
(+82) 010-4584-3490
woojae1123@gmail.com / wkim97@kaist.ac.kr
Personal Webpage
Github *Google Scholar*

Education

2023 – present: **Ph.D., School of Computing**, *KAIST, Daejeon, Korea.*
Advisor: Sung-Eui Yoon
Research topics: AI robustness and safety

2021 – 2023: **M.S., School of Computing**, *KAIST, Daejeon, Korea.*
Advisor: Sung-Eui Yoon
Research topics: Adversarial attack
Thesis: Diverse Generative Perturbations on Attention Space for Transferable Adversarial Attacks

2016 – 2021: **B.S., School of Computing**, *KAIST, Daejeon, Korea.*
Minor in Electrical Engineering

2012 – 2016: **High school**, *Northview High School, Johns Creek, GA, USA.*

Research Interests

Core Focus: Adversarial robustness, AI safety, and 3D vision.

First-Authored Contributions:.

• **3D vision security:** Introduced AegisRF (BMVC'25), sensitivity-guided adversarial perturbations for protecting intellectual property of neural radiance fields from unauthorized use
• **Multi-threat robustness:** Developed RoME (under review, CVPR'26), a mixture-of-experts framework with LoRA adapters achieving robustness across diverse adversarial threats
• **Single-threat robustness:** Created FSR (CVPR'23 Highlights, $\sim$ 2.6% acceptance), a feature separation approach that recalibrates non-robust feature maps for enhanced robustness
• **Transferable attacks:** Designed ADA (ICIP'22 Oral), generating diverse perturbations on attention space to achieve high transferability across different model architectures

Collaborative Research:.

Extensive contributions to **3D computer vision** (Gaussian splatting, neural radiance fields, inverse rendering, sparse-view surface reconstruction), **AI safety and privacy** (membership inference, diffusion-based manipulation protection), person re-identification, and video processing

Publications

First-authored

[C.15] RoME: Robust Mixture of LoRA Experts against Multiple Adversarial Perturbations.
Woo Jae Kim, Kyle Min, Suhyeon Ha, Joonsung Jeon, and Sung-Eui Yoon.
CVPR, 2026 (*under review*)

[C.12] AegisRF: Adversarial Perturbations Guided with Sensitivity for Protecting Intellectual Property of Neural Radiance Fields.
Woo Jae Kim, Kyu Beom Han, Youngju Na, Yoonki Cho, Junsik Jung, Sooel Son, and Sung-Eui Yoon.
BMVC, 2025
[Paper] [Code]

[C.5] Feature Separation and Recalibration for Adversarial Robustness.
Woo Jae Kim, Yoonki Cho, Junsik Jung, and Sung-Eui Yoon.
***CVPR**, 2023*
Highlights paper ($\sim$ 2.6% acceptance rate)
[Paper] [Code]

[C.3] Diverse Generative Perturbations on Attention Space for Transferable Adversarial Attacks.
Woo Jae Kim, Seunghoon Hong, and Sung-Eui Yoon.
***ICIP**, 2022*
Oral paper ($\sim$ 10% acceptance rate)
[Paper] [Code]

Co-authored

[C.14] No Caption, No Problem: Caption-Free Membership Inference via Model-Fitted Embeddings.
Joonsung Jeon, **Woo Jae Kim**, Suhyeon Ha, Sooel Son, and Sung-Eui Yoon.
ICLR, 2026 (*under review*)

[C.13] Radiometrically Consistent Gaussian Surfels for Inverse Rendering.
Kyu Beom Han, Jaeyoon Kim, **Woo Jae Kim**, Jinhwan Seo, and Sung-Eui Yoon.
ICLR, 2026 (*under review*)

[C.11] Learning Event-guided Exposure-agnostic Video Frame Interpolation via Adaptive Feature Blending.
Junsik Jung, Yoonki Cho, **Woo Jae Kim**, Lin Wang, and Sung-Eui Yoon.
[Paper]
BMVC, 2025

[C.10] Pose-free 3D Gaussian splatting via shape-ray estimation.
Youngju Na, Taeyeon Kim, Jumin Lee, Kyu Beom Han, **Woo Jae Kim**, and Sung-Eui Yoon.
***ICIP**, 2025*
Best student paper award (1 out of 643 papers)
[Paper]

[C.9] AdvPaint: Protecting Images from Inpainting Manipulation via Adversarial Attention Disruption.
Joonsung Jeon, **Woo Jae Kim**, Suhyeon Ha, Sooel Son, and Sung-eui Yoon.
***ICLR**, 2025*
[Paper] [Code]

[C.8] Generalizable Person Re-identification via Balancing Alignment and Uniformity.
Yoonki Cho, Jaeyoon Kim, **Woo Jae Kim**, Junsik Jung, and Sung-eui Yoon.
***NeurIPS**, 2024*
[Paper] [Code]

[C.7] UFORecon: Generalizable Sparse-View Surface Reconstruction from Arbitrary and Unfavorable Data Pairs.
Youngju Na, **Woo Jae Kim**, Kyu Beom Han, Suhyeon Ha, and Sung-Eui Yoon.
CVPR, *2024*
[Paper] [Code]

[C.6] Towards Content-based Pixel Retrieval in Revisited Oxford and Paris.
Guoyuan An, **Woo Jae Kim**, Saelyne Yang, Rong Li, Yuchi Huo, and Sung-Eui Yoon.
ICCV, *2023*
[Paper] [Code]

[C.4] Pixel-wise Guidance for Utilizing Auxiliary Features in Monte Carlo Denoising.
Kyubeom Han, Olivia G. Odenthal, **Woo Jae Kim**, and Sung-Eui Yoon.
i3D, *2023*
[Paper] [Code]

[C.2] Part-based Pseudo Label Refinement for Unsupervised Person Re-identification.
Yoonki Cho, **Woo Jae Kim**, Seunghoon Hong, and Sung-Eui Yoon.
CVPR, *2022*
[Paper] [Code]

[C.1] Deep Video Inpainting Guided by Audio-Visual Self-Supervision.
Kyuyeon Kim, Junsik Jung*, **Woo Jae Kim***, and Sung-Eui Yoon. (* *equal contributions*)
ICASSP, *2022*
[Paper] [Code]

Fellowships & Awards

Nov. 2023 ***Recipient*** of the Qualcomm Innovation Fellowship.

Feb. 2022 ***Paper Award*** in the 34th Workshop on Image Processing and Image Understanding (IPIU).

Aug. 2023, Feb. 2022 ***Best TA Award*** in School of Computing, KAIST.

Feb. 2021 ***Grand Prix Award (1st place)*** in the Undergraduate Research Program (URP) in KAIST.

Invited Talks & Presentations

Jan. 2024 Invited to Qualcomm Korea to give a talk on adversarial robustness [C.5].

Feb. 2022 Invited to Korean Conference on Computer Vision (KCCV) 2023 for oral and poster presentations on adversarial robustness [C.5].

Skills

Languages: Korean Native, English Native

Programming Languages: Python, C, C++, Java, MATLAB, R

Tools: PyTorch, Tensorflow, Keras, Numpy, LaTex, Kubernetes, Docker

Web Technologies: HTML, CSS, Javascript

Professional Service

2023-2025 **CVPR**, Reviewer.

2024 **ECCV**, Reviewer.

2023,2025 **ICCV**, Reviewer.

2025 **NeurIPS**, Reviewer.

Teaching Experience

2024 **Teaching Assistant, KAIST**.
CS588 Deep Learning based Image Search

2023, 2025 **Teaching Assistant, Samsung Electronics**.
AI Expert Program

2018-2019, 2021-2023 **Teaching Assistant, KAIST**.
CS101 Introduction to Programming

2021, 2023 **Teaching Assistant, KAIST**.
CS206 Data Structure