

Sangmin Woo

PH.D. CANDIDATE IN EE @ KAIST

291, Daehak-ro, Yuseong-gu, Daejeon, 34141, Rep. of KOREA

✉ smwoo95@kaist.ac.kr | 🏠 sangminwoo.github.io | 📄 sangminwoo | 🎓 [Sangmin Woo](https://sangminwoo)

I am currently pursuing a Ph.D. degree in Electrical Engineering at [KAIST](https://kaist.ac.kr). In 2021, I completed an M.S. degree in Electrical Engineering and Computer Science at [GIST](https://gist.ac.kr). Prior to that, I obtained a B.S. degree in Electrical Engineering from [KNU](https://knu.ac.kr) in 2019.

I thrive on creative challenges and enjoy building strong relationships along the way. Explore my academic journey below, and contact me directly to learn more.

Research Interest

Humans are inherently **multi-modal** learners, naturally understanding the world by looking (vision), listening (audio), and communicating (language). I am passionate about advancing machine intelligence to mirror this ability, enabling systems to interpret the world holistically and **generate** creative, human-centered content.

My work explores the following, but not limited to:

- **Multi-modal AI**: Vision + {Language, Audio, etc.}
- **Generative AI** (Diffusion Models, Vision Language Models)
- **Image** / **Video Understanding**

Research Experience

Amazon AWS AI

RESEARCH INTERN

Remote

Sep. 2024 - Mar. 2025

Amazon AWS AI

RESEARCH INTERN

Santa Clara, CA, United States

Jun. 2024 - Sep. 2024

NAVER LABS

RESEARCH INTERN

Suwon, Korea

Apr. 2023 - Aug. 2023

Publication

2025

Black-Box Visual Prompt Engineering for Mitigating Object Hallucination in Large Vision Language Models

2025

NATIONS OF THE AMERICAS CHAPTER OF THE ASSOCIATION FOR COMPUTATIONAL LINGUISTICS (**NAACL MAIN**)

Gen AI & Multi-modal

[Paper](#)

[Sangmin Woo](#), Kang Zhou, Yun Zhou, Shuai Wang, Sheng Guan, Haibo Ding, Lin Lee Cheong

Diffusion Model Patching via Mixture-of-Prompts

2025

AAAI CONFERENCE ON ARTIFICIAL INTELLIGENCE (**AAAI**)

Gen AI

[Paper](#) | [Code](#) | [Project](#)

Seokil Ham*, [Sangmin Woo*](#), Jinyoung Kim, Hyojun Go, Byeongjun Park, Changick Kim (*: Equal Contribution)

2024

Parameter Efficient Mamba Tuning via Projector-targeted Diagonal-centric Linear Transformation

ARXIV

[Paper](#)

Seokil Ham, Hee-Seon Kim, [Sangmin Woo](#), Changick Kim

2024

Learning

RITUAL: Random Image Transformations as a Universal Anti-hallucination Lever in LVLMS

ARXIV

[Paper](#) | [Code](#) | [Project](#)

[Sangmin Woo*](#), Jaehyuk Jang*, Donguk Kim*, Yubin Choi, Changick Kim (*: Equal Contribution)

2024

Gen AI & Multi-modal

Don't Miss the Forest for the Trees: Attentional Vision Calibration for Large Vision Language Models

ARXIV

[Paper](#) | [Code](#) | [Project](#)

[Sangmin Woo*](#), Donguk Kim*, Jaehyuk Jang*, Yubin Choi, Changick Kim (*: Equal Contribution)

2024

Gen AI & Multi-modal

Flow-Assisted Motion Learning Network for Weakly-Supervised Group Activity Recognition

EUROPEAN CONFERENCE ON COMPUTER VISION (ECCV)

[Paper](#)

Muhammad Adi Nugroho, [Sangmin Woo](#), Sumin Lee, Jinyoung Park, Yooseung Wang, Donguk Kim, Changick Kim

2024

Multi-modal & video

Spatio-Temporal Proximity-Aware Dual-Path Model for Panoramic Activity Recognition

EUROPEAN CONFERENCE ON COMPUTER VISION (ECCV)

[Paper](#)

Sumin Lee, Yooseung Wang, [Sangmin Woo](#), Changick Kim

2024

video

Switch Diffusion Transformer: Synergizing Denoising Tasks with Sparse Mixture-of-Experts

EUROPEAN CONFERENCE ON COMPUTER VISION (ECCV)

[Paper](#) | [Code](#) | [Project](#)

Byeongjun Park, Hyojun Go, Jinyoung Kim, [Sangmin Woo](#), Seokil Ham*, Changick Kim

2024

Gen AI

HarmonyView: Harmonizing Consistency and Diversity in One-Image-to-3D

COMPUTER VISION AND PATTERN RECOGNITION CONFERENCE (CVPR)

[Paper](#) | [Code](#) | [Project](#) | [Demo](#)

[Sangmin Woo*](#), Byeongjun Park*, Hyojun Go, Jinyoung Kim, Changick Kim (*: Equal Contribution)

2024

Gen AI

Denoising Task Routing for Diffusion Models

INTERNATIONAL CONFERENCE OF LEARNING REPRESENTATION (ICLR)

[Paper](#) | [Code](#) | [Project](#)

Byeongjun Park*, [Sangmin Woo*](#), Hyojun Go*, Jinyoung Kim*, Changick Kim (*: Equal Contribution)

2024

Gen AI

Sketch-based Video Object Localization

WINTER CONFERENCE ON APPLICATIONS OF COMPUTER VISION (WACV)

[Paper](#) | [Code](#)

[Sangmin Woo](#), Soyeong Jeon, Jinyoung Park, Minji Son, Sumin Lee, Changick Kim

2024

Multi-modal & video

2023

AHFu-Net: Align, Hallucinate, and Fuse Network for Missing Multimodal Action Recognition

INTERNATIONAL CONFERENCE ON VISUAL COMMUNICATIONS AND IMAGE PROCESSING (VCIP) (ORAL PRESENTATION)

Muhammad Adi Nugroho, [Sangmin Woo](#), Sumin Lee, Changick Kim

2023

Multi-modal & video

Multi-modal Social Group Activity Recognition in Panoramic Scene

INTERNATIONAL CONFERENCE ON VISUAL COMMUNICATIONS AND IMAGE PROCESSING (VCIP)

Donguk Kim, Sumin Lee, [Sangmin Woo](#), Jinyoung Park, Muhammad Adi Nugroho, Changick Kim

2023

Multi-modal & video

Cross-Modal Alignment and Translation for Missing Modality Action Recognition

COMPUTER VISION AND IMAGE UNDERSTANDING (CVIU)

[Paper](#)

Yeonju Park, [Sangmin Woo](#), Sumin Lee, Muhammad Adi Nugroho, Changick Kim

2023

Multi-modal & video

Modality Mixer Exploiting Complementary Information for Multi-modal Action Recognition

COMPUTER VISION AND IMAGE UNDERSTANDING (CVIU) – MAJOR REVISION

[Paper](#)

Sumin Lee, [Sangmin Woo](#), Yeonju Park, Muhammad Adi Nugroho, Changick Kim

2023

Multi-modal & video

Audio-Visual Glance Network for Efficient Video Recognition

INTERNATIONAL CONFERENCE ON COMPUTER VISION (ICCV)

[Paper](#)

Muhammad Adi Nugroho, [Sangmin Woo](#), Sumin Lee, Changick Kim

2023

Multi-modal & video

Towards Good Practices for Missing Modality Robust Action Recognition

AAAI CONFERENCE ON ARTIFICIAL INTELLIGENCE (AAAI) (ORAL PRESENTATION)

[Paper](#) | [Code](#)

[Sangmin Woo](#), Sumin Lee, Yeonju Park, Muhammad Adi Nugroho, Changick Kim

2023

Multi-modal & video

Modality Mixer for Multi-modal Action Recognition

WINTER CONFERENCE ON APPLICATIONS OF COMPUTER VISION (WACV)

[Paper](#)

Sumin Lee, [Sangmin Woo](#), Yeonju Park, Muhammad Adi Nugroho, Changick Kim

2023

Multi-modal & video

~2022

Explore-And-Match: Bridging Proposal-Based and Proposal-Free with Transformer for Sentence Grounding in Videos

ARXIV

[Paper](#) | [Code](#)

[Sangmin Woo](#), Jinyoung Park, Inyong Koo, Sumin Lee, Minki Jeong, Changick Kim

2022

Multi-modal & video

Tackling the Challenges in Scene Graph Generation with Local-to-Global Interactions

TRANSACTIONS ON NEURAL NETWORKS AND LEARNING SYSTEMS (TNNLS)

[Paper](#) | [Code](#)

[Sangmin Woo](#), Junhyug Noh, Kangil Kim

2022

Multi-modal & Image

Temporal Flow Mask Attention for Open-Set Long-Tailed Recognition of Wild Animals in Camera-Trap Images

INTERNATIONAL CONFERENCE ON IMAGE PROCESSING (ICIP)

[Paper](#)

Jeongsoo Kim, [Sangmin Woo](#), Byeongjun Park, Changick Kim

2022

Image

Impact of Sentence Representation Matching in Neural Machine Translation

APPLIED SCIENCES

[Paper](#)

Heeseung Jung, Kangil Kim, Jong-Hun Shin, Seung-Hoon Na, SangKeun Jung, [Sangmin Woo](#)

2022

Learning

What and When to Look?: Temporal Span Proposal Network for Video Relation Detection

EXPERT SYSTEMS WITH APPLICATIONS (ESWA) – MAJOR REVISION

[Paper](#) | [Code](#)

[Sangmin Woo](#), Junhyug Noh, Kangil Kim

2021

video

Revisiting Dropout: Escaping Pressure for Training Neural Networks with Multiple Costs

ELECTRONICS

[Paper](#) | [Code](#)

[Sangmin Woo](#), Kangil Kim, Junhyug Noh, Jong-Hun Shin, Seung-Hoon Na

2021

Learning

KOREAN DOMESTIC

Mitigating Hallucinations in Vision Language Models with High-pass filter

INSTITUTE OF ELECTRONICS AND INFORMATION ENGINEERS (IEIE)

Donguk Kim, Sangmin Woo, Jaehyuk Jang, Yubin Choi, Changick Kim

2024

Gen AI & Multi-modal

Contrastive Decoding with Image Processing Filters for Mitigating Object Hallucinations in Large-scale Vision-Language Models

INSTITUTE OF ELECTRONICS AND INFORMATION ENGINEERS (IEIE)

Yubin Choi, Donguk Kim, Sangmin Woo, Jaehyuk Jang

2024

Gen AI & Multi-modal

Light-Weighted Korean Speech Recognition System for Edge Devices

INSTITUTE OF ELECTRONICS AND INFORMATION ENGINEERS (IEIE)

Yooseung Wang, Sangmin Woo, Changick Kim

2023

Learning

On Learning Relations between Objects in Images

KOREA INSTITUTE OF MILITARY SERVICE AND TECHNOLOGY (KIMST)

Sangmin Woo, Changick Kim

2022

Multi-modal & Image

Effective Trash Classification using Attentional Learning

KOREA SOFTWARE CONGRESS (KSC)

[Code](#)

Sangmin Woo, Soon Ki Jung

2018

Image

Honors & Awards

| | |
|-----------|---|
| Dec, 2024 | Finalist , Qualcomm Innovation Fellowship 2024 Korea |
| Oct, 2023 | Invited Paper Talk , Center for Applied Research in Artificial Intelligence (CARAI) Workshop |
| Dec, 2022 | Finalist , 29th HumanTech Paper Award @ Samsung Electronics Co., Ltd. |
| Dec, 2021 | Top Award (\$ 10,000) , LG Electronics Robot Contest @ LG Electronics Co., Ltd. |
| Nov, 2019 | Excellence Award (\$ 500) , Creative Space G A.I&IoT Makerthon @ GIST |

Academic Activity

I serve as a reviewer in the following conferences and journals.

COMPUTER VISION

| | |
|--|--------|
| IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR) | 2024 ~ |
| European Conference on Computer Vision (ECCV) | 2024 ~ |

MACHINE LEARNING

| | |
|---|--------|
| Annual Conference on Neural Information Processing Systems (NeurIPS) | 2024 ~ |
| International Conference on Learning Representations (ICLR) | 2024 ~ |
| International Conference on Machine Learning (ICML) | 2025 ~ |
| AAAI Conference on Artificial Intelligence (AAAI) | 2023 ~ |
| International Conference on Artificial Intelligence and Statistics (AISTATS) | 2025 ~ |

JOURNALS

IEEE Transactions on Neural Networks and Learning Systems (TNNLS)
IEEE Transactions on Circuits and Systems for Video Technology (TCSVT)

Patent

Group Activity Recognition Apparatus and Method using RGB Videos and LiDAR Data

KR PATENT 10-2741168

Changick Kim, Jinyoung Park, Donguk Kim, Sumin Lee, Muhammad Adi Nugroho, Sangmin Woo, Yooseung Wang

2024

Method and Device for Inferring Dynamic Relationship between Objects in Video 2023

KR PATENT 10-2590916

Sangmin Woo, Kangil Kim

Human Activity Recognition Apparatus, Human Activity Recognition Method and Computer Program for Classifying Activity Categories of User 2023

KR PATENT: 10-2515065

Changick Kim, Inyong Koo, Yeonju Park, Minki Jeong, Sumin Lee, Sangmin Woo

Scene Graph Generation Apparatus 2021

KR PATENT 10-2254768

Sangmin Woo, Kangil Kim

Project

Scene Text Recognition with Visual Contexts 2024.02 2024.12

CENTER FOR SECURITY TECHNOLOGY RESEARCH, KAIST

Multi-modal Group Activity Recognition 2023.02 2024.02

CENTER FOR APPLIED RESEARCH IN ARTIFICIAL INTELLIGENCE (CARAI)

Sketch-based Video Object Localization 2023.02 2023.11

CENTER FOR SECURITY TECHNOLOGY RESEARCH, KAIST

Multi-modal Action Recognition 2021.09 2022.12

CENTER FOR APPLIED RESEARCH IN ARTIFICIAL INTELLIGENCE (CARAI)

Development of Precise Content Identification Technology based on Relationship Analysis for Maritime Vessels/Structure 2021.09 2021.12

MINISTRY OF SCIENCE AND ICT (MSIT)

Education

Korea Advanced Institute of Science and Technology (KAIST) Daejeon, Korea

PH.D. IN ELECTRICAL ENGINEERING

Aug. 2021 - Aug. 2025

Gwangju Institute of Science and Technology (GIST) Gwangju, Korea

M.S. IN ELECTRICAL ENGINEERING AND COMPUTER SCIENCE

Sep. 2019 - Aug. 2021

Kyungpook National University (KNU) Daegu, Korea

B.S. IN ELECTRICAL ENGINEERING

Mar. 2013 - Aug. 2019