Sangmin **Woo**

APPLIED SCIENTIST @ AMAZON

2795 Augustine Dr, Santa Clara, CA 95054, United States

✉ sangminw@amazon.com | 🏠 sangminwoo.github.io | ⬜ sangminwoo | 🎓 Sangmin Woo

I am an Applied Scientist at Amazon Agentic AI. I earned my Ph.D. from KAIST, during which I did research internships at Amazon and Naver Labs. 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, Large Vision Language Models, Large Language Models)
- Image / Video Understanding

Research Experience

Amazon Agentic AI	*Santa Clara, CA, United States*
APPLIED SCIENTIST	*Jun. 2025 - Present*
Amazon Bedrock	*Remote*
APPLIED SCIENTIST INTERN	*Sep. 2024 - Mar. 2025*
Amazon Bedrock	*Santa Clara, CA, United States*
APPLIED SCIENTIST INTERN	*Jun. 2024 - Sep. 2024*
NAVER LABS	*Suwon, Korea*
RESEARCH INTERN	*Apr. 2023 - Aug. 2023*

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*

Honors & Awards

Jun, 2025	**Outstanding Reviewer (711/12,593),** IEEE Conference on Computer Vision and Pattern Recognition (CVPR)
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 ~
International Conference on Computer Vision (ICCV) 2025 ~
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 ~

JOURNALS

IEEE Transactions on Neural Networks and Learning Systems (TNNLS)
Transactions on Machine Learning Research (TMLR)

Publication

2025

Don't Miss the Forest for the Trees: Attentional Vision Calibration for Large Vision Language Models *2025*

FINDINGS OF THE ASSOCIATION FOR COMPUTATIONAL LINGUISTICS (**ACL FINDINGS**) *Gen AI* & *Multi-modal*

Paper | Code | Project

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

Modality Mixer Exploiting Complementary Information for Multi-modal Action Recognition *2025*

COMPUTER VISION AND IMAGE UNDERSTANDING (**CVIU**) *Multi-modal* & *video*

Paper

Sumin Lee, Sangmin Woo, Yeonju Park, Muhammad Adi Nugroho, Changick Kim

Parameter Efficient Mamba Tuning via Projector-targeted Diagonal-centric Linear Transformation *2025*

COMPUTER VISION AND PATTERN RECOGNITION CONFERENCE (**CVPR**) *Learning*

Paper

Seokil Ham, Hee-Seon Kim, Sangmin Woo, Changick Kim

A Systematic Survey of Automatic Prompt Optimization Techniques *2025*

ARXIV *Gen AI*

Paper

Kiran Ramnath, Kang Zhou, Sheng Guan, Soumya Smruti Mishra, Xuan Qi, Zhengyuan Shen, Shuai Wang, Sangmin Woo, Sullam Jeoung, Yawei Wang, Haozhu Wang, Han Ding, Yuzhe Lu, Zhichao Xu, Yun Zhou, Balasubramaniam Srinivasan, Qiaojing Yan, Yueyan Chen, Haibo Ding, Panpan Xu, Lin Lee Cheong

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**)

`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

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

2024

ARXIV

`Gen AI` & `Multi-modal`

Paper | Code | Project

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

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

2024

EUROPEAN CONFERENCE ON COMPUTER VISION (**ECCV**)

`Multi-modal` & `video`

Paper

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

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

2024

EUROPEAN CONFERENCE ON COMPUTER VISION (**ECCV**)

`video`

Paper

Sumin Lee, Yooseung Wang, Sangmin Woo, Changick Kim

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

2024

EUROPEAN CONFERENCE ON COMPUTER VISION (**ECCV**)

`Gen AI`

Paper | Code | Project

Byeongjun Park, Hyojun Go, Jinyoung Kim, Sangmin Woo, Seokil Ham*, Changick Kim

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

2024

COMPUTER VISION AND PATTERN RECOGNITION CONFERENCE (**CVPR**)

`Gen AI`

Paper | Code | Project | Demo

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

Denoising Task Routing for Diffusion Models

2024

INTERNATIONAL CONFERENCE OF LEARNING REPRESENTATION (**ICLR**)

`Gen AI`

Paper | Code | Project

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

Sketch-based Video Object Localization

2024

WINTER CONFERENCE ON APPLICATIONS OF COMPUTER VISION (**WACV**)

`Multi-modal` & `video`

Paper | Code

Sangmin Woo, Soyeong Jeon, Jinyoung Park, Minji Son, Sumin Lee, Changick Kim

2023

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

2023

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

`Multi-modal` & `video`

Muhammad Adi Nugroho, Sangmin Woo, Sumin Lee, Changick Kim

Multi-modal Social Group Activity Recognition in Panoramic Scene

2023

INTERNATIONAL CONFERENCE ON VISUAL COMMUNICATIONS AND IMAGE PROCESSING (**VCIP**)

`Multi-modal` & `video`

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

Cross-Modal Alignment and Translation for Missing Modality Action Recognition

Computer Vision and Image Understanding (**CVIU**)

2023

Multi-modal & *video*

Paper

Yeonju Park, <u>Sangmin Woo</u>, Sumin Lee, Muhammad Adi Nugroho, Changick Kim

Audio-Visual Glance Network for Efficient Video Recognition

International Conference on Computer Vision (**ICCV**)

2023

Multi-modal & *video*

Paper

Muhammad Adi Nugroho, <u>Sangmin Woo</u>, Sumin Lee, Changick Kim

Towards Good Practices for Missing Modality Robust Action Recognition

AAAI Conference on Artificial Intelligence (**AAAI**) (**Oral Presentation**)

2023

Multi-modal & *video*

Paper | Code

<u>Sangmin Woo</u>, Sumin Lee, Yeonju Park, Muhammad Adi Nugroho, Changick Kim

Modality Mixer for Multi-modal Action Recognition

Winter Conference on Applications of Computer Vision (**WACV**)

2023

Multi-modal & *video*

Paper

Sumin Lee, <u>Sangmin Woo</u>, Yeonju Park, Muhammad Adi Nugroho, Changick Kim

~2022

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

Arxiv

2022

Multi-modal & *video*

Paper | Code

<u>Sangmin Woo</u>, Jinyoung Park, Inyong Koo, Sumin Lee, Minki Jeong, Changick Kim

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

Transactions on Neural Networks and Learning Systems (**TNNLS**)

2022

Multi-modal & *Image*

Paper | Code

<u>Sangmin Woo</u>, Junhyug Noh, Kangil Kim

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

International Conference on Image Processing (**ICIP**)

2022

Image

Paper

Jeongsoo Kim, <u>Sangmin Woo</u>, Byeongjun Park, Changick Kim

Impact of Sentence Representation Matching in Neural Machine Translation

Applied Sciences

2022

Learning

Paper

Heeseung Jung, Kangil Kim, Jong-Hun Shin, Seung-Hoon Na, SangKeun Jung, <u>Sangmin Woo</u>

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

Expert Systems With Applications (**ESWA**) – Major Revision

2021

video

Paper | Code

<u>Sangmin Woo</u>, Junhyug Noh, Kangil Kim

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

Electronics

2021

Learning

Paper | Code

<u>Sangmin Woo</u>, Kangil Kim, Junhyug Noh, Jong-Hun Shin, Seung-Hoon Na

Korean Domestic

Mitigating Hallucinations in Vision Language Models with High-pass filter

Institute of Electronics and Information Engineers (**IEIE**)

2024

Gen AI & *Multi-modal*

Donguk Kim, <u>Sangmin Woo</u>, Jaehyuk Jang, Yubin Choi, Changick Kim

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

2024

INSTITUTE OF ELECTRONICS AND INFORMATION ENGINEERS (**IEIE**)

`Gen AI` & `Multi-modal`

Yubin Choi, Donguk Kim, Sangmin Woo, Jaehyuk Jang

Light-Weighted Korean Speech Recognition System for Edge Devices

2023

INSTITUTE OF ELECTRONICS AND INFORMATION ENGINEERS (**IEIE**)

`Learning`

Yooseung Wang, Sangmin Woo, Changick Kim

On Learning Relations between Objects in Images

2022

KOREA INSTITUTE OF MILITARY SERVICE AND TECHNOLOGY (**KIMST**)

`Multi-modal` & `Image`

Sangmin Woo, Changick Kim

Effective Trash Classification using Attentional Learning

2018

KOREA SOFTWARE CONGRESS (**KSC**)

`Image`

Code

Sangmin Woo, Soon Ki Jung

Patent

Black-Box Visual Prompt Engineering for Improving Large Vision Language Models Across Visual Tasks

2025

US PATENT

Sangmin Woo, Kang Zhou, Yun Zhou, Shuai Wang, Sheng Guan, Haibo Ding, Lin Lee Cheong

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

2024

KR PATENT 10-2741168

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

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