## Sangmin Woo

**INCOMING APPLIED SCIENTIST @ AMAZON** 

2795 Augustine Dr, Santa Clara, CA 95054, US

🔽 sangminw@amazon.com 📔 🖀 sangminwoo.github.io 📔 🖸 sangminwoo 📔 🞓 Sangmin Woo

I will be joining Amazon Agentic AI as an Applied Scientist in June 2025. I received my Ph.D. degree at KAIST. I have industry experience through internships at Amazon AWS 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

## **Res**earch Experience \_\_\_\_

## AWS Agentic AI

Applied Scientist

AWS Bedrock

Applied Scientist Intern

#### AWS Bedrock

Applied Scientist Intern

#### NAVER LABS

Research Intern

## **Publication**

#### 2025

Don't Miss the Forest for the Trees: Attentional Vision Calibration for Large Vision	2025
Language Models	2025
Findings of the Association for Computational Linguistics ( <b>ACL Findings</b> )	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	2025
Recognition	2025
Computer Vision and Image Understanding (CVIU)	Multi-modal & video
Paper	

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

Jun. 2025 - Present

Santa Clara, CA, United States

Remote Sep. 2024 - Mar. 2025

## Santa Clara, CA, United States

Jun. 2024 - Sep. 2024

Suwon, Korea Apr. 2023 - Aug. 2023

Parameter Efficient Mamba Tuning via Projector-targeted Diagonal-centric Linear	2025
<b>Transformation</b> Computer Vision and Pattern Recognition Conference ( <b>CVPR</b> )	Learning
Paper	
Seokil Ham, Hee-Seon Kim, <u>Sangmin Woo</u> , Changick Kim	
A Systematic Survey of Automatic Prompt Optimization Techniques	2025
Arxiv Paper	Gen Al
Kiran Ramnath, Kang Zhou, Sheng Guan, Soumya Smruti Mishra, Xuan Qi, Zhengyuan Shen, Shuai Wang, <u>Sangmin Woo</u> Wang, Haozhu Wang, Han Ding, Yuzhe Lu, Zhichao Xu, Yun Zhou, Balasubramaniam Srinivasan, Qiaojing Yan, Yueyan Ch Xu, Lin Lee Cheong	, Sullam Jeoung, Yawei ien, Haibo Ding, Panpan
Black-Box Visual Prompt Engineering for Mitigating Object Hallucination in Large Vision	2025
Language Models	Con AL & Multi model
Paper	Gen Ar & Multi-modul
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 Al
Paper   Code   Project Seokil Ham* Sangmin Woo* Jinvoung Kim Hyoiun Go Byeongiun Park Changick Kim (* Equal Contribution)	
2024	
RITUAL: Random Image Transformations as a Universal Anti-hallucination Lever in LVLMs	2024
Arxiv	Gen Al & Multi-modal
Paper   Code   Project	
Sangmin Woo", Jaenyuk Jang", Donguk Kim", Yubin Choi, Changick Kim (": Equal Contribution)	
Flow-Assisted Motion Learning Network for Weakly-Supervised Group Activity Perognition	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 ( <b>ECCV</b> ) Paper	video
Sumin Lee, Yooseung Wang, <u>Sangmin Woo</u> , Changick Kim	
Switch Diffusion Transformer: Synergizing Denoising Tasks with Sparse	
Mixture-of-Experts	2024
EUROPEAN CONFERENCE ON COMPUTER VISION (ECCV)	Gen Al
Paper   Code   Project Byeongjun Park, Hyojun Go, Jinyoung Kim, Sangmin Woo, Seokil Ham*, Changick Kim	
Harmony/liow: Harmonizing Consistancy and Diversity in One-Image-to 2D	2024
Computer Vision and Pattern Recognition Conference ( <b>CVPR</b> )	Gen Al
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) Paper   Code   Project	Gen Al
Byeongjun Park*, <u>Sangmin Woo*</u> , 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 ( <b>VCIP</b> ) (Oral Presentation) Muhammad Adi Nugroho, <u>Sangmin Woo</u> , Sumin Lee, Changick Kim	Multi-modal & video
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, <u>Sangmin Woo</u> , Jinyoung Park, Muhammad Adi Nugroho, Changick Kim	
Cross-Modal Alignment and Translation for Missing Modality Action Recognition	2023
Computer Vision and Image Understanding ( <b>CVIU</b> ) Pader	Multi-modal & video
Yeonju Park, Sangmin Woo, Sumin Lee, Muhammad Adi Nugroho, Changick Kim	
Audio-Visual Glance Network for Efficient Video Recognition	2023
International Conference on Computer Vision (ICCV)	Multi-modal & video
Paper	
Munammad Adi Nugrono, <u>Sangmin Woo</u> , Sumin Lee, Changick Kim	
Towards Good Practices for Missing Modality Robust Action Recognition	2023
Paper   Code	Mulli-modal & video
Sangmin Woo, Sumin Lee, Yeonju Park, Muhammad Adi Nugroho, Changick Kim	
Modality Mixer for Multi-modal Action Recognition	2023
Winter Conference on Applications of Computer Vision (WACV)	Multi-modal & video
Paper	
Sumin Lee, <u>Sangmin Woo</u> , Yeonju Park, Munammad Adi Nugrono, Changick Kim	
~2022	
Explore-And-Match: Bridging Proposal-Based and Proposal-Free with Transformer for	2022
Sentence Grounding in Videos	2022
Arxiv Paper   Code	Multi-modal & video
Sangmin Woo, Jinyoung Park, Inyong Koo, Sumin Lee, Minki Jeong, Changick Kim	
Tackling the Challenges in Scene Graph Generation with Local-to-Global Interactions	2022
Transactions on Neural Networks and Learning Systems (TNNLS)	Multi-modal & Image
Paper   Code	
Sangmin Woo, Junhyug Noh, Kangil Kim	
Temporal Flow Mask Attention for Open-Set Long-Tailed Recognition of Wild Animals in	2022
Camera-Trap Images	Imago
Paper	innage
Jeongsoo Kim, <u>Sangmin Woo</u> , Byeongjun Park, Changick Kim	
Impact of Sentence Representation Matching in Neural Machine Translation	2022
Applied Sciences	Learning
Paper	

Heeseung Jung, Kangil Kim, Jong-Hun Shin, Seung-Hoon Na, SangKeun Jung, Sangmin Woo

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	
<b>Mitigating Hallucinations in Vision Language Models with High-pass filter</b> INSTITUTE OF ELECTRONICS AND INFORMATION ENGINEERS (IEIE) Donguk Kim, <u>Sangmin Woo</u> , Jaehyuk Jang, Yubin Choi, Changick Kim	2024 Gen Al & 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 Al & 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 <mark>Image</mark>

## 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~
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 ~

International Conference on Artificial Intelligence and Statistics (AISTATS) 2025 ~

## JOURNALS

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

## Patent\_\_\_\_\_

Black-Box Visual Prompt Engineering for Improving Large Vision Language Models Across Visual Tasks US PATENT Sangmin Woo, Kang Zhou, Yun Zhou, Shuai Wang, Sheng Guan, Haibo Ding, Lin Lee Cheong		2025
Group Activity Recognition Apparatus and Method using RGB Videos and LiDAR Data KR Ратемт 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 KR Patent 10-2590916 Sangmin Woo, Kangil Kim		2023
Human Activity Recognition Apparatus, Human Activity Recognition Method and Computer Program for Classifying Activity Categories of User KR PATENT: 10-2515065 Changick Kim, Inyong Koo, Yeonju Park, Minki Jeong, Sumin Lee, Sangmin Woo		2023
Scene Graph Generation Apparatus KR PATENT 10-2254768 Sangmin Woo, Kangil Kim		2021
Project		
Scene Text Recognition with Visual Contexts Center for Security Technology Research, KAIST	2024.02	2024.12
Multi-modal Group Activity Recognition Center for Applied Research in Artificial Intelligence (CARAI)	2023.02	2024.02
Sketch-based Video Object Localization Center for Security Technology Research, KAIST	2023.02	2023.11
Multi-modal Action Recognition Center for Applied Research in Artificial Intelligence (CARAI)	2021.09	2022.12
Development of Precise Content Identification Technology based on Relationship Analysis for Maritime Vessels/Structure Ministry of Science and ICT (MSIT)	2021.09	2021.12

Education \_\_\_\_\_

## Korea Advanced Institute of Science and Technology (KAIST)

Ph.D. IN ELECTRICAL ENGINEERING

Daejeon, Korea Aug. 2021 - Aug. 2025

## Gwangju Institute of Science and Technology (GIST)

M.S. IN ELECTRICAL ENGINEERING AND COMPUTER SCIENCE

## Kyungpook National University (KNU)

B.S. IN ELECTRICAL ENGINEERING

**Gwangju, Korea** Sep. 2019 - Aug. 2021

**Daegu, Korea** Mar. 2013 - Aug. 2019