



# SANGHO LEE

✉ sanghol@allenai.org | 🏠 <https://sangho-vision.github.io> |  sanghol |  Scholar

## RESEARCH INTERESTS

---

Computer Vision, Machine Learning  
Multimodal representation learning, especially for high-level video understanding and reasoning

## EMPLOYMENT

---

**PRIOR @ Allen Institute for AI** *Aug. 2024 - Present*  
Research Scientist

**PRIOR @ Allen Institute for AI** *Jan. 2023 - Jul. 2024*  
Postdoctoral Researcher (Young Investigator)

**PRIOR @ Allen Institue for AI** *Mar. 2022 - Jun. 2022, Aug. 2022 - Dec. 2022*  
Research Intern

## EDUCATION

---

**Seoul National University** *Mar. 2017 - Feb. 2023*  
Doctor of Philosophy (Ph.D) Overall GPA: 4.27 / 4.3  
Department of Computer Science and Engineering  
Advisor: Prof. Gunhee Kim  
Thesis: Improving Efficiency in Large-Scale Self-Supervised Video Representation Learning

**Seoul National University** *Mar. 2010 - Feb. 2017*  
Bachelor of Science Overall GPA: 4.04 / 4.3  
Department of Computer Science and Engineering  
Minor in Statistics  
Graudated *summa cum laude*

## PUBLICATIONS

---

### **Molmo and PixMo: Open Weights and Open Data for State-of-the-Art Multimodal Models**

Matt Deitke\*, Christopher Clark\*, **Sangho Lee**, Rohun Tripathi, Yue Yang, Jae Sung Park, Mohammadreza Salehi, Niklas Muennighoff, Kyle Lo, Luca Soldaini, Jiasen Lu, Taira Anderson, Erin Bransom, Kiana Ehsani, Huong Ngo, YenSung Chen, Ajay Patel, Mark Yatskar, Chris Callison-Burch, Andrew Head, Rose Hendrix, Favyen Bastani, Eli VanderBilt, Nathan Lambert, Yvonne Chou, Arnavi Chheda, Jenna Sparks, Sam Skjonsberg, Michael Schmitz, Aaron Sarnat, Byron Bischoff, Pete Walsh, Chris Newell, Piper Wolters, Tanmay Gupta, Kuo-Hao Zeng, Jon Borchardt, Dirk Groeneveld, Jen Dumas, Crystal Nam, Sophie Lebrecht, Caitlin Wittlif, Carissa Schoenick, Oscar Michel, Ranjay Krishna, Luca Weihs, Noah A. Smith, Hannaneh Hajishirzi, Ross Girshick, Ali Farhadi, and Aniruddha Kembhavi (\*: equal contribution)

Conference on Computer Vision and Pattern Recognition 2025 (CVPR 2025), Oral

🏆 **Best Paper, Honorable Mention**

### **One Diffusion to Generate Them All**

Duong H. Le\*, Tuan Pham\*, **Sangho Lee**, Christopher Clark, Aniruddha Kembhavi, Stephan Mandt, Ranjay Krishna, and Jiasen Lu (\*: equal contribution)

Conference on Computer Vision and Pattern Recognition 2025 (CVPR 2025)

### **ReSpec: Relevance and Specificity Grounded Online Filtering for Learning on Video-Text Data Streams**

Chris Dongjoo Kim\*, Jihwan Moon\*, Sangwoo Moon, Heeseung Yun, Sihaeng Lee, Aniruddha Kembhavi, Soonyoung Lee, Gunhee Kim, **Sangho Lee**<sup>^</sup>, and Christopher Clark<sup>^</sup> (\*: equal contribution, <sup>^</sup>: equal advising)

Conference on Computer Vision and Pattern Recognition 2025 (CVPR 2025)

**Finding NeMo: Negative-mined Mosaic Augmentation for Referring Image Segmentation**

Seongsu Ha\*, Chaeyun Kim\*, Donghwa Kim\*, Junho Lee, **Sangho Lee**, and Joonseok Lee (\*: equal contribution)

The 18th European Conference on Computer Vision (ECCV 2024)

**Unified-IO 2: Scaling Autoregressive Multimodal Models with Vision, Language, Audio, and Action**

Jiasen Lu\*, Christopher Clark\*, **Sangho Lee**\*, Zichen Zhang\*, Savya Khosla, Ryan Marten, Derek Hoiem, and Aniruddha Kembhavi (\*: equal contribution)

Conference on Computer Vision and Pattern Recognition 2024 (CVPR 2024), Highlight

**Can Language Models Laugh at YouTube Short-form Videos?**

Dayoon Ko, **Sangho Lee**, and Gunhee Kim

The 2023 Conference on Empirical Methods in Natural Language Processing (EMNLP 2023)

**ACAV100M: Automatic Curation of Large-Scale Datasets for Audio-Visual Video Representation Learning**

**Sangho Lee**\*, Jiwan Chung\*, Youngjae Yu, Gunhee Kim, Thomas Breuel, Gal Chechik, and Yale Song (\*: equal contribution)

International Conference on Computer Vision 2021 (ICCV 2021)

CVPR 2021: The Third Workshop on Learning from Unlabeled Videos

**Unsupervised Representation Learning via Neural Activation Coding**

Yookoon Park, **Sangho Lee**, Gunhee Kim, and David Blei

The Thirty-eighth International Conference on Machine Learning (ICML 2021), Long Talk

**Parameter Efficient Multimodal Transformers for Video Representation Learning**

**Sangho Lee**, Youngjae Yu, Gunhee Kim, Thomas Breuel, Jan Kautz, and Yale Song

The Ninth International Conference on Learning Representations (ICLR 2021)

CVPR 2021: The Second International Workshop on Large Scale Holistic Video Understanding

**Self-Supervised Learning of Compressed Video Representations**

Youngjae Yu\*, **Sangho Lee**\*, Gunhee Kim, and Yale Song (\*: equal contribution)

The Ninth International Conference on Learning Representations (ICLR 2021)

**A Memory Network Approach for Story-based Temporal Summarization of 360° Videos**

**Sangho Lee**, Jinyoung Sung, Youngjae Yu, and Gunhee Kim

Conference on Computer Vision and Pattern Recognition 2018 (CVPR 2018)

ECCV 2018 Workshop on 360° Perception and Interaction

**A Deep Ranking Model for Spatio-temporal Highlight Detection from a 360° Video**

Youngjae Yu, **Sangho Lee**, Joonil Na, Jaeyoun Kang, and Gunhee Kim

The Thirty-Second AAAI Conference on Artificial Intelligence (AAAI-18), Spotlight

**A Read-Write Memory Network for Movie Story Understanding**

Seil Na, **Sangho Lee**, Jisung Kim, and Gunhee Kim

International Conference on Computer Vision 2017 (ICCV 2017)

ICCV 2017: The Joint Video and Language Understanding Workshop

**Encoding Video and Label Priors for Multi-label Video Classification on YouTube-8M dataset**

Seil Na, Youngjae Yu, **Sangho Lee**, Jisung Kim, and Gunhee Kim  
CVPR 2017 Workshop on YouTube-8M Large-Scale Video Understanding

## PROFESSIONAL ACTIVITIES

---

### Workshop Organization and Committee

Workshop on on Video-Language Models at NeurIPS 2024 (1st)

### Journal Review

Transactions on Machine Learning Research (TMLR) 2025/2024/2023

### Conference Review

CVPR 2024/2022/2021

ICCV 2023

ECCV 2024/2022

NeurIPS 2024/2023

ICML 2024

ICLR 2025/2024/2023/2021

AAAI 2023

### Workshop Paper Review

CVPR Workshop on Computer Vision in the Wild 2025

NeurIPS Workshop on Self-Supervised Learning – Theory and Practice 2024/2023/2022

## AWARDS

---

### Best Paper Awards @ CVPR 2025

*Jun. 2025*

Molmo and PixMo: Open Weights and Open Data for State-of-the-Art Multimodal Models

Best Paper, Honorable Mention

### Excellent Ph.D. Thesis Award

*Feb. 2023*

Selected as the best doctoral thesis by Department of Computer Science and Engineering, Seoul National University

### Naver Ph.D. Fellowship

*Dec. 2021*

Awarded to outstanding graduate students in the field of Computer Science for their exceptional academic research

### Youlchon AI Star Fellowship

*Sep. 2021*

An award for those who made distinguished research achievements in core AI fields

### MovieQA Challenge @ ICCV 2017 Workshop

*Oct. 2017*

ICCV 2017 Workshop on the Joint Video and Language Understanding Workshop

Ranked 2nd place

### Google Cloud & YouTube-8M Video Understanding Challenge

*Jul. 2017*

CVPR 2017 Workshop on YouTube-8M Large-Scale Video Understanding

Ranked 8th place out of 655 teams (Top 2%)

## ACADEMIC EXPERIENCE

---

### Graduate Teaching Assistant at SNU

Knowledge Representation and Reasoning (M1522.001300)

Fall 2018

Probabilistic Graphical Models (M1522.001300)

Fall 2017

Discrete Mathematics (4190.101)

Spring 2017