

Heeseung Kwon

POSITION	Post-doctoral Researcher / Research Engineer KAIST ALIN Lab / RLWRLD
CONTACT INFORMATION	RLWRLD Rubina Building B1, 561 Seolleung-ro, Gangnam-gu, Seoul, South Korea e-mail: heeseung.kwon@rlwrl.dai
CITIZENSHIP	The Republic of Korea
BIRTHDAY	July 28th, 1989.
RESEARCH INTERESTS	Computer Vision & Machine Learning, especially, Vision-Language-Action Models.
EDUCATION	Pohang University of Science and Technology (POSTECH) , Pohang, Korea <i>Ph.D., (Unified master's and doctor's course)</i> Sep 2012 – Feb 2021 • Thesis: Learning Temporal Dynamics for Video Action Recognition • Advisor: Prof. Minsu Cho • Co-Advisor: Prof. Suha Kwak Pohang University of Science and Technology (POSTECH) , Pohang, Korea <i>B.S., Department of Electronic Engineering</i> Mar 2008 – Jun 2012
RESEARCH EXPERIENCE	THOTH research team , INRIA Grenoble-Rhône-Alpes, France <i>Post-doctoral Researcher</i> Sep 2021 – Dec 2024 • Under the supervision of Kartee Alahari. Dep. of Computer Science and Engineering , POSTECH, Korea <i>Post-doctoral Researcher</i> Feb 2021 – Aug 2021 • Research on video understanding and dynamic neural networks. • Collaboration with students at Computer Vision Lab. (Manjin Kim, Dahyun Kang, Juhong Min).
INTERNATIONAL PUBLICATIONS	Chunghyun Park*, Beomjun Kim* Seungcheol Park [†] Heeseung Kwon[†] , Yashu Shukla, Seunghoon Sim, Jinwoo Shin, Minsu Cho, “PointWAM: 3D World Action Modeling for Dexterous Robotic Manipulation”, in ICLR 2027 , (* co-first authors, [†] equal contribution) (under review). Manjin Kim*, Heeseung Kwon* , Kartee Alahari, Minsu Cho, “Exploring Multi-order Self-Similarity for Motion Understanding”, in NeurIPS 2026 (*equal contribution). Dongyoung Kim ^{†§} , Huiwon Jang ^{†§} , Myungkyu Koo [†] , Suhyeok Jang [†] , Taeyoung Kim [†] , Beomjun Kim, Byungjun Yoon, Changsung Jang, Daewon Choi, Dongsu Han, Donguk Lee, Heeseung Kwon , Hojin Jeon, Jaehyun Kang, Jaekyoung Bae, Jihyuk Lee, Jimin Lee, John Won, Joonwoo Ahn, Junhyeong Park, Junyoung Sung, Kyungmin Lee, Minseong Han, Minsung Yoon, Sejune Joo, Seonil Son, Seungcheol Park, Seunggeun Cho, Seungjun Moon, Seungku Kim, Yonghoon Dong, Yongjin Cho, Youngchan Kim, Jinwoo Shin [§] , “RLDX-1 Technical Report”, in arXiv 2026 ([†] project leads, [§] research leads).

Huiwon Jang, Sihyun Yu, **Heeseung Kwon**, Hojin Jeon, Younggyo Seo*, and Jinwoo Shin*, “ContextVLA: Vision-Language-Action Model with Amortized Multi-frame Context”, in **ICLR**, 2025 (*equal contribution, under review).

Manjin Kim, **Heeseung Kwon**, Karteek Alahari and Minsu Cho, ”Exploring High-order Self-Similarity for Video Understanding”, in **arXiv**, 2025.

Heeseung Kwon, Francisco M. Castro, Manuel J Marin-Jimenez, Nicolas Guil and Karteek Alahari, “Lightweight Structure-Aware Attention for Visual Understanding.”, in **IJCV**, 2025.

Manjin Kim*, **Heeseung Kwon***, Chunyu Wang, Suha Kwak and Minsu Cho, “Relational Self-Attention: What’s Missing in Attention for Video Understanding.”, in **NeurIPS**, 2021 (*equal contribution).

Dahyun Kang, **Heeseung Kwon**, Juhong Min and Minsu Cho, “Correlational Embedding for Few-Shot Classification.”, in **ICCV**, 2021.

Heeseung Kwon*, Manjin Kim*, Suha Kwak and Minsu Cho, “Learning Self-Similarity in Space and Time as Generalized Motion for Video Action Recognition.”, in **ICCV**, 2021 (*equal contribution).

Gyeongsik Moon*, **Heeseung Kwon***, Kyoung Mu Lee and Minsu Cho, “IntegralAction: Pose-driven Feature Integration for Robust Human Action Recognition in Videos.”, in the 2nd International Workshop on Large-scale Holistic Video Understanding, **CVPR Workshop**, 2021 (*equal contribution).

Heeseung Kwon, Manjin Kim, Suha Kwak, and Minsu Cho, “MotionSqueeze: Neural Motion Feature Learning for Video Understanding” in **ECCV**, 2020.

Heeseung Kwon, Woohyun Shim, and Minsu Cho, “Temporal U-Nets for Video Summarization with Scene and Action Recognition”, in CoVieW’19 Proceedings of the 2nd Workshop and Challenge on Comprehensive Video Understanding in the Wild, **ICCV Workshop**, 2019 (Challenge winner as 2nd place).

Heeseung Kwon, Suha Kwak, and Minsu Cho, “Video Understanding via Convolutional Temporal Pooling Network and Multimodal Feature Fusion”, CoVieW’18 Proceedings of the 1st Workshop and Challenge on Comprehensive Video Understanding in the Wild, ACM Multimedia conference (**MM Workshop**), 2018 (Challenge winner as 1st place).

Heeseung Kwon, Yeonho Kim, Jin S. Lee, and Minsu Cho, “First Person Action Recognition via Two-stream ConvNet with Long-term Fusion Pooling” in Pattern Recognition Letters (**PRL**), Vol. 68, no. 1, pp. 76-82, 2018.

Seungjae Oh, **Heeseung Kwon**, and Hyojeong So, “Hidden UI: projection-based augmented reality for map navigation on multi-touch tabletop”, Proceedings of the 2nd ACM symposium on Spatial User Interaction (**SUI**), 2014.

HONOURS AND
AWARDS

The challenge winner at [CoVieW’19 workshop and challenge](#), ICCV 2019

- 2nd place at the Challenge on Comprehensive Video Understanding in the Wild

The challenge winner at [CoVieW’18 workshop and challenge](#), ACM MM 2018

- 1st place at the Challenge on Comprehensive Video Understanding in the Wild
- Awarded to the best algorithm on multi-task action and scene recognition

ACADEMIC SERVICES	I served as a reviewer in CVPR, ECCV, CVIU, IJCV.	
PROJECTS	Development of advanced face recognition system integration co-op with POSCO ICT	Mar 2020 – Dec 2020
	Source technology research for context recognition and prediction via comprehensive video understanding co-op with Hanyang University and KAIST	May 2018 – Dec 2020
	Development of object detection and semantic segmentation system for intelligent vehicles co-op with VADAS	Oct 2018 – Dec 2019
	Development of the multi-view multi-channel broadcasting system for sports digital contents funded by Ministry of Culture, Sports and Tourism	Jun 2016 – Jun 2018
	Enhancing usability of 2D screen and projection based augmented reality system funded by Department of Creative IT Engineering	Jan 2014 – Dec 2015
PATENTS	Yeongsuk Kim, Heeseung Kwon , “Method for temporal information encoding of the video segment frame-wise features for video recognition” (Korean domestic; 1019369470000)	
LANGUAGE SKILLS	Korean(native), English(advanced)	
PROGRAMMING	Python, C, C++, MATLAB, Pytorch, Keras, Tensorflow, L ^A T _E X	
REFEREES	<i>Available on request.</i>	