Dahun Kim

Senior Research Scientist, Google DeepMind

Google Gradient Canopy, Mountain View, CA 94043
mcahny@google.com | https://mcahny.github.io

Summary

- **Research Interest:** Vision-and-Language, Computer Vision, Machine Learning – My research interests lie in understanding the interaction of vision and language, and enabling new capabilities with multimodal models.
- **Research Experience:** Prior to Google DeepMind, I was a Research Scientist at Google Brain. I also worked at Adobe Research, Google Research and KAIST.

Research Experience

• **Google DeepMind (previously Google Brain)**, MTV, CA — Jul.2022 - Present
Senior Research Scientist, Vision-and-Language team

• **Google Research**, MTV, CA — Jul.2022 - Apr.2023
Research Intern, worked on "video mask transformer"

• **Google Brain**, MTV, CA — Jun.2020 - Nov.2020
Research Intern, worked on "open-world detection - detect everything"

• **Adobe Research**, San Jose, CA — Jun.2019 - Sep.2019
Research Intern, worked on "video panoptic segmentation"

Education

• **Ph.D.** in Dept. of **Electrical Eng., KAIST** — Mar.2018 - Feb.2022
Advised by Prof. In So Kweon
Thesis: "Learning Dense Pixel Features for Video Processing and Understanding"
Best Thesis Award from KAIST EE.

• **M.S.** in Dept. of **Electrical Eng., KAIST** — Mar.2016 - Feb.2018
Advised by Prof. In So Kweon
Thesis: "Reducing Human Supervision in Supervised Learning"

• **B.S.** in Dept. of **Electrical Eng., KAIST** — Mar.2012 - Feb.2016

Academic Service

- **Area Chair** in ICML 2025
- **Area Chair** in NeurIPS 2025, 2024, 2023
- **Area Chair** in CVPR 2025, 2024, 2023
- Journal Reviewer in TPAMI, TNNLS, TIP, EuroGraphics
- Conference Reviewer in CVPR [22, 21, 20], NeurIPS [21, 20], ECCV [24, 20], ICCV [25, 23, 21, 19], ICLR [24, 21], AAAI [24, 22, 21, 20],

Publications

038. **Gemini Team**
"**Gemini 2.5**: Pushing the Frontier with Advanced Reasoning, Multimodality, Long Context, and Next Generation Agentic Capabilities"

037. **Dahun Kim**, A. Angelova
"Context-Adaptive Multi-Prompt Embedding with Large Language Models for Vision-Language Alignment"
in **COLM 2025**, Montreal, Canada

036. **Dahun Kim**, A. Piergiovanni, G. Mallya, A. Angelova
"VideoComp: Advancing Fine-Grained Compositional and Temporal Alignment in Video-Text Models"
in **CVPR 2025**, Nashville, USA

035. S. Wang, **Dahun Kim**, A. Taalimi, C. Sun, W. Kuo
"Learning Visual Grounding from Generative Vision and Language Model"
in **WACV 2025**, Arizona, USA

034. A. Piergiovanni **Dahun Kim**, M. Ryoo, I. Noble, A. Angelova
"Whats in a Video: Factorized Autoregressive Decoding for Online Dense Video Captioning"
in **Preprint**

033. **Dahun Kim**, A. Angelova, W. Kuo
"Region-centric Image-Language Pretraining for Open-Vocabulary Detection"
in **ECCV 2024**, Milano, Italy
Launched at Google Vertex AI model garden

032. M. Kim, J. Choi, **Dahun Kim**, Y. M. Ro
"Textless Unit-to-Unit training for Many-to-Many Multilingual Speech-to-Speech Translation"
in **TASLP 2024**, *IEEE Trans. on Audio, Speech and Language Processing*

031. A. Piergiovanni, I. Nobel, **Dahun Kim**, M. Ryoo, V. Gomes, A. Angelova
"Mirasol3B: A Multimodal Autoregressive model for time-aligned and contextual modalities"
in **CVPR 2024**, Seattle, USA
Featured in Google AI BlogPost

031. J. Kim, H. Oh, B. Kwon, **Dahun Kim**, Y. Kwon, T. Oh
"Uni-DVPS: Unified Model for Depth-Aware Video Panoptic Segmentation"
in **RA-L** and **ICRA 2024 (Oral)**; *IEEE Robotics and Automation Letters* ,

030. **Dahun Kim**, A. Angelova, W. Kuo
"Contrastive Feature Masking Open-vocabulary Vision Transformer",
in **ICCV 2023**, Paris, France

029. **Dahun Kim**, A. Angelova, W. Kuo
"Region-Aware Pretraining for Open-Vocabulary Object Detection with Vision Transformers",
in **CVPR 2023 (Highlight - accept rate: 2.5%)**, Vancouver, Canada
Featured in Google AI BlogPost

028. W. Kuo†, A. Piergiovanni†, **Dahun Kim***, X. Luo*, B. Caine, W. Li, A. Ogale, L. Zhou, A. Dai, Z. Chen, C. Cui, A. Angelova
"MaMMUT: A Simple Vision-Encoder Text-Decoder Architecture for Multimodal Tasks",
in **TMLR 2023**: Transactions on Machine Learning Research
Featured in Google AI BlogPost

027. R. Li, **Dahun Kim**, W. Kuo
"RECLIP: Resource-efficient CLIP by Training with Small Images",
in **TMLR 2023**: Transactions on Machine Learning Research

026. Shin, **Dahun Kim**, Q. Yu, J. Xie, H.S. Kim, B. Green, I.S. Kweon, K.J. Yoon, L.C. Chen
"Video-kMaX: A Simple Unified Approach for Online and Near-Online Video Panoptic Segmentation",
in **WACV 2024 (Oral)** and **CVPRW 2023**: 'Transformers for Vision' Workshop

025. Y. Kwon, **Dahun Kim**, D. Ceylan, H. Fuchs
"Neural Image-based Avatars: Generalizable Radiance Fields for Human Avatar Modeling",
in **ICLR 2023**, Kigali, Rwanda

024. **Dahun Kim**, S. Woo, J.Y. Lee, I.S. Kweon
"Dense Pixel-level Interpretation of Dynamic Scenes with Video Panoptic Segmentation",
in **TIP 2022**: *IEEE Trans. on Image Processing*, IF=10.6

023. **Dahun Kim**, J. Xie, H. Wang, S. Qiao, H.S. Kim, H. Adam, I.S. Kweon, L.C. Chen
"TubeFormer-DeepLab: video mask transformer",
in **CVPR 2022**, New Orleans, USA
Ranked #1 on SemKITTI-DVPS, #3 on KITTI-STEP benchmark

022. Q. Yu, H. Wang, **Dahun Kim**, S. Qiao, M. Collins, Y. Zhu, H. Adam, A. Yuille, L.C. Chen
"CMT-DeepLab: dynamic clustering mask transformers for panoptic segmentation",
in **CVPR 2022 (Oral)**, New Orleans, USA

021. **Dahun Kim**, T.Y. Lin, A. Angelova, I. S. Kweon, W. Kuo
"Learning open-world object proposals without learning to classify",
in **RA-L** and **ICRA 2022 (Oral)**; *IEEE Robotics and Automation Letters*, Philadelphia, USA
Invited paper talk at Open-World Segmentation (UVO) Workshop @ ICCV 2021
Received Qualcomm Innovation Award 2021

020. Y. Kwon, S. Petrangeli, **Dahun Kim**, H. Wang, V. Swaminathan, H. Fuchs
"Tailor Me: An Editing Network for Fashion Attribute Shape Manipulation".
in **WACV 2022 (Oral)**

019. Y. Kwon, **Dahun Kim**, D. Ceylan, H. Fuchs
"Neural Human Performer: learning generalizable radiance fields for human performance rendering",
in **NeurIPS 2021 (Spotlight)**, Virtual
Received Bronze Prize, 28th Samsung HumanTech Paper Award
018. S. Woo, **Dahun Kim**, J.Y. Lee, I. S. Kweon,
"Learning to associate every segment for video panoptic segmentation".
in **CVPR 2021**, Virtual

017. S. Woo, **Dahun Kim**, J.Y. Lee, I.S. Kweon
"Global Context and Geometric Priors for Effective Non-Local Self-Attention".
in **BMVC 2021**
Received Bronze Prize, 27th Samsung HumanTech Paper Award

016. M. Kim, S. Woo, **Dahun Kim**, I. S. Kweon,
"The Devil is in the Boundary: Exploiting Boundary Representation for Basis-based Instance Segmentation".
in **WACV 2021 (Oral)**

015. Y. Kwon, S. Petrangeli, **Dahun Kim**, H. Wang, V. Swaminathan, H. Fuchs,
"Rotationally-Temporally Consistent Novel View Synthesis for Human Performance Video",
in **ECCV 2020 (Spotlight)**, Virtual

014. **Dahun Kim**, S. Woo, J.Y. Lee, I.S. Kweon,
"Video panoptic segmentation",
in **CVPR 2020 (Oral)**, Virtual

013. **Dahun Kim***, S. Woo*, J.Y. Lee, I.S. Kweon,
"Recurrent temporal aggregation framework for deep video inpainting",
in **TPAMI 2020**: *IEEE Trans. on Pattern Analysis and Machine Intelligence*, IF=24.314
Received KAIST-Samsung Industry-University Cooperation Best Paper Award

012. Y. Jung, **Dahun Kim**, S. Woo, K. Kim, S. Kim, I.S. Kweon,
"Hide-and-Tell: Learning to bridge photo streams for visual storytelling",
in **AAAI 2020**, New York, USA (Acceptance: 1591/7737 ≈ 20.6%)

011. Y. Kwon, S. Petrangeli, **Dahun Kim**, H. Wang, H. Fuchs, V. Swaminathan,
"Rotationally-Consistent Novel View Synthesis for Humans",
in ACM **MM 2020**, Virtual (Acceptance: 472/1698 ≈ 27.8%)

010. S. Woo, **Dahun Kim**, K. Park, J.Y. Lee, I.S. Kweon,
"Align-and-Attend Network for Globally and Locally Coherent Video Inpainting",
in **BMVC 2020** (Acceptance: 195/670 ≈ 29.1%)

009. **Dahun Kim***, S. Woo*, J.Y. Lee, I.S. Kweon,
"Deep video inpainting",
in **CVPR 2019**, Long Beach, USA (Acceptance: 1294/5160 ≈ 25.2%)

008. **Dahun Kim***, S. Woo*, J.Y. Lee, I.S. Kweon,
"Deep blind video decaptioning by temporal aggregation and recurrence",
in **CVPR 2019**, Long Beach, USA (Acceptance: 1294/5160 ≈ 25.2%)
1st place winner of ECCV 2018 Chalearn LAP Video De-Captioning Challenge

007. **Dahun Kim**, D. Cho, I.S. Kweon,
"Self-supervised video representation learning with space-time cubic puzzles",
in **AAAI 2019 (Oral)**, Honolulu, USA

006. Y. Jung, D. Cho, **Dahun Kim**, S. Woo, I.S. Kweon,
"Discriminative feature learning for unsupervised video summarization",
in **AAAI 2019 (Oral)**, Honolulu, USA
Received Honorable Mention, 25th Samsung HumanTech Paper Award

005. K. Park, S. Woo, **Dahun Kim**, D. Cho, I.S. Kweon,
"Preserving Semantic and Temporal Consistency for Unpaired Video-to-Video Translation",
in ACM **MM 2019**, Nice, France (Acceptance: 252/936 ≈ 26.9%)

004. Cho, Y. Jung, F. Rameau, **Dahun Kim**, S. Woo, I.S. Kweon,
"Video Retargeting: Trade-off between Content Preservation and Spatio-temporal Consistency",
in ACM **MM 2019**, Nice, France (Acceptance: 252/936 ≈ 26.9%)

003. S. Woo*, **Dahun Kim***, D. Cho, I.S. Kweon,
"LinkNet: relational embedding for scene graph",
in **NeurIPS 2018**, Montreal, Canada (Acceptance: 1011/4856 ≈ 20.8%)

002. **Dahun Kim**, D. Cho, D. Yoo, I.S. Kweon,
"Learning image representations by completing damaged jigsaw puzzles",
in **WACV 2018 (Oral)**, Lake Tahoe, USA

001. **Dahun Kim**, D. Cho, D. Yoo, I.S. Kweon,
"Two-phase learning for weakly supervised object localization",
in **ICCV 2017**, Venice, Italy (Acceptance: 621/2143 ≈ 28.9%)

Patents

Issued:

P1. Video Panoptic Segmentation (US Patent 11,640,714)

P2. Panoptic Segmentation (US Patent 11,256,960)

Pending:

P3. Method and Device for Hierarchical Learning of Neural Network Based on Weakly Supervised Learning (US Patent App. 16/758,089)

P4. Electronic Device and Control Method of Same (US Patent App. 17/554,142)

P5. Methods and apparatus localizing object (s) in vision data (US Patent App. 18/289,725)

Awards and Honors

- Best Ph.D. Thesis Award, EE, KAIST — Apr.2022
- Bronze Award, 28th HumanTech Paper Award, Samsung Electronics Co., Ltd. ($5,000) — Feb.2022
- Qualcomm Innovation Award (Korea) 2021 — Nov.2021
- Outstanding Reviewers Award, CVPR 2021 — Aug.2021
- Outstanding Reviewers Award, ECCV 2020 — Aug.2020
- Microsoft Research Asia (MSRA) Ph.D Fellowship 2019 Winner ($10,000) — Oct.2019
- 1-st Place Award in ChaLearnLAP 2018 Inpainting Challenge Track 2: video decaptioning (ECCV2018 Challenge) — Sep.2018
- Global Ph.D Fellowship, National Research Foundation of Korea (National Minister fellowship – ≈ $60,000 + 3-year full scholarship) — Mar.2018 - Feb.2021
- KAIST-Samsung Industry-University Cooperation, Best Paper Award ($3,000) — Jul.2020
- Bronze Award, 27th HumanTech Paper Award, Samsung Electronics Co., Ltd. ($5,000) — Feb.2021
- Honorable Mention, 25th HumanTech Paper Award, Samsung Electronics Co., Ltd. ($2,000) — Feb.2019
- Lab Student Representative (over 30 members), — Sep.2019 - Jul.2020
- Bronze Prize, Best Paper Award, 31th IPIU — Feb.2019
- International Computer Vision Summer School (ICVSS), Sicily, Italy — Jul.2018

References

Prof. In So Kweon: M.S. - Ph.D. advisor at KAIST
KEPCO Chair Professor, Dept. of EE, KAIST; Eemail: iskweon77@kaist.ac.kr

Dr. Anelia Angelova: Collaborator at Google
Principal Research Scientist, Google DeepMind; Email: anelia@google.com

Dr. Weicheng Kuo: Collaborator at Google
Staff Research Scientist, Google DeepMind; Email: weicheng@google.com

Dr. Liang-Chieh Chen: Collaborator at Google
Senior Principal Scientist, Amazon; Email: lcchen@cs.ucla.edu

Dr. Joon-Young Lee: Collaborator at Adobe
Manager and Senior Research Scientist, Adobe Research; Email: jolee@adobe.com

Dr. Tsung-Yi Lin: Collaborator at Google
Principal Research Scientist, Nvidia; Email: tsungyilin87@gmail.com