

Jaemin Cho

PHD STUDENT @ UNC CHAPEL HILL

[✉ jmincho@cs.unc.edu](mailto:jmincho@cs.unc.edu) | [🌐 j-min.io](http://j-min.io) | [🐙 j-min](https://github.com/j-min) | [🌐 jaemin-cho](https://www.linkedin.com/in/jaemin-cho) | [📷 Jaemin Cho](https://www.instagram.com/jaemincho)

Interests

Machine Learning, Multimodal Learning, Natural Language Processing, Generative Models

Education

University of North Carolina at Chapel Hill

PH.D. IN COMPUTER SCIENCE (ADVISOR: MOHIT BANSAL)

Chapel Hill, NC

Aug. 2020 - Present

University of North Carolina at Chapel Hill

M.S. IN COMPUTER SCIENCE (ADVISOR: MOHIT BANSAL)

Chapel Hill, NC

Aug. 2020 - May 2022

Seoul National University

B.S. IN INDUSTRIAL ENGINEERING

Seoul, Korea

Mar. 2011 - Feb. 2018.

Research Experience

Google Research

STUDENT RESEARCHER (ADVISOR: SU WANG, JORDI PONT-TUSET, JASON BALDRIDGE)

Austin, TX

May. 2023 - Aug. 2023

- Worked on Text-to-Image Evaluation

Microsoft Research

RESEARCH INTERN (ADVISOR: LINJIE LI, ZHENGYUAN YANG, ZHE GAN, LIJUAN WANG)

Redmond, WA (Remote)

May. 2022 - Aug. 2022

- Worked on Layout-to-Image Generation

Adobe Research

RESEARCH INTERN (ADVISOR: DAVID SEUNGHYUN YOON, AJINKYA KALE, FRANCK DERNONCOURT, TRUNG BUI)

San Jose, CA (Remote)

May. 2021 - Aug. 2021

- Worked on Fine-grained Image Captioning

Allen Institute for AI

PREDOCTORAL YOUNG INVESTIGATOR (ADVISOR: JIASEN LU, HANNANEH HAJISHIRZI, ANI KEMBHAVI)

Seattle, WA

Sep. 2019 - Jul. 2020

- Worked on Text-to-Image Generation with Multimodal Transformers

SNU Music & Audio Research Group

VISITING SCHOLAR (ADVISOR: KYOGU LEE)

Seoul, Korea

Jun. 2019 - Jul. 2019.

- Worked on 3D Pose Estimation, Music & Dance Modeling

Naver Clova

AI RESIDENCY (ADVISOR: MINJOON SEO, HANNANEH HAJISHIRZI)

Seongnam, Korea

Jul. 2018 - Mar. 2019

- Worked on Two-stage Diverse Text Generation

SNU Vision & Learning Lab

RESEARCH INTERN (ADVISOR: GUNHEE KIM)

Seoul, Korea

Mar. 2017 - May 2018

- Worked on Hierarchical VAE for Conversation Modeling

HKUST

RESEARCH INTERN (ADVISOR: SUNGHUN KIM)

Hong Kong

Jan. 2017 - Feb. 2017

- Worked on Abstract Syntax Tree-based Python-Pseudocode Translation

Polytechnique Montréal

RESEARCH INTERN (ADVISOR: NADIA LAHRICHI)

Montreal, Canada

Jan. 2015 - Jun. 2015

- Worked on Patient Classification & Nursing Scheduling based on Mixed-Integer Programming

Industry Experience

DataNada

NLP ENGINEER (MANAGER: JOHN PARK)

Seoul, Korea

Aug. 2016 - Dec. 2016

- Worked on Dialog Management & Dependency Parsing for Chatbot

Teaching Experience

Pre-doctoral Student Mentoring at UNC

- Zineng Tang: UG at UNC (2021 - 2022) [[Homepage](#)]
- Abhay Zala: UG/MS at UNC (2021 - Present) [[Homepage](#)]

Teaching Assistant at SNU

- 4190.666 Machine Learning (Fall 2017) (Instructor: Hyun Oh Song) [[Link](#)]

Instructor at FastCampus

- Tutorial on Deep Learning for Chatbot Development (Sep 2017) [[Course materials](#)]

Publications

Conference Papers

- [C18] Qin Liu, **Jaemin Cho**, Mohit Bansal, and Marc Niethammer. “Rethinking Interactive Image Segmentation with Low Latency, High Quality, and Diverse Prompts”. *CVPR*. 2024. [[Project](#)] [[Paper](#)] [[Code](#)].
- [C17] **Jaemin Cho**, Yushi Hu, Roopal Garg, Peter Anderson, Ranjay Krishna, Jason Baldridge, Mohit Bansal, Jordi Pont-Tuset, and Su Wang. “Davidsonian Scene Graph: Improving Reliability in Fine-Grained Evaluation for Text-to-Image Generation”. *ICLR*. 2024. [[Project](#)] [[Paper](#)] [[Code](#)].
- [C16] **Jaemin Cho**, Abhay Zala, and Mohit Bansal. “Visual Programming for Text-to-Image Generation and Evaluation”. *NeurIPS*. 2023. [[Project](#)] [[Paper](#)] [[VPGen Code](#)] [[VPEval Code](#)].
- [C15] Shoubin Yu, **Jaemin Cho**, Prateek Yadav, and Mohit Bansal. “Self-Chaind Image-Language Model for Video Localization and Question Answering”. *NeurIPS*. 2023. [[Paper](#)] [[Code](#)].
- [C14] Zhenhailong Wang, Ansel Blume, Sha Li, Genglin Liu, **Jaemin Cho**, Zineng Tang, Mohit Bansal, and Heng Ji. “Paxion: Patching Action Knowledge in Video-Language Foundation Models”. *NeurIPS (spotlight)*. 2023. [[Paper](#)] [[Code](#)].
- [C13] **Jaemin Cho**, Abhay Zala, and Mohit Bansal. “DALL-Eval: Probing the Reasoning Skills and Social Biases of Text-to-Image Generation Models”. *ICCV*. 2023. [[Paper](#)] [[Code](#)].
- [C12] Abhay Zala*, **Jaemin Cho***, Satwik Kottur, Xilun Chen, Barlas Oğuz, Yashar Mehdad, and Mohit Bansal. “Hierarchical Video-Moment Retrieval and Step-Captioning”. *CVPR*. 2023. [[Project](#)] [[Paper](#)] [[Code](#)].
- [C11] Zineng Tang*, **Jaemin Cho***, Jie Lei, and Mohit Bansal. “Perceiver-VL: Efficient Vision-and-Language Modeling with Iterative Latent Attention”. *WACV*. 2023. [[Paper](#)] [[Code](#)].
- [C10] Zineng Tang*, **Jaemin Cho***, Yixin Nie*, and Mohit Bansal. “Textless Vision-Language Transformer”. *NeurIPS (oral)*. 2022. [[Paper](#)] [[Code](#)].
- [C9] Yi-Lin Sung, **Jaemin Cho**, and Mohit Bansal. “LST: Ladder Side-Tuning for Parameter and Memory Efficient Transfer Learning”. *NeurIPS*. 2022. [[Paper](#)] [[Code](#)].
- [C8] **Jaemin Cho**, Seunghyun Yoon, Ajinkya Kale, Franck Dernoncourt, Trung Bui, and Mohit Bansal. “Fine-grained Image Captioning with CLIP Reward”. *Findings of NAACL (short)*. 2022. [[Paper](#)] [[Code](#)].
- [C7] Yi-Lin Sung, **Jaemin Cho**, and Mohit Bansal. “VL-Adapter: Parameter-Efficient Transfer Learning for Vision-and-Language Tasks”. *CVPR*. 2022. [[Paper](#)] [[Code](#)].
- [C6] Revanth Gangi Reddy, Xilin Rui, Manling Li, Xudong Lin, Haoyang Wen, **Jaemin Cho**, Lifu Huang, Mohit Bansal, Avi Sil, Shih-Fu Chang, Alexander Schwing, and Heng Ji. “MuMuQA: Multimedia Multi-Hop News Question Answering via Cross-Media Knowledge Extraction and Grounding”. *AAAI*. 2022. [[Paper](#)].
- [C5] Zineng Tang, **Jaemin Cho**, Hao Tan, and Mohit Bansal. “VidLanKD: Improving Language Understanding via Video-Distilled Knowledge Transfer”. *NeurIPS*. 2021. [[Paper](#)] [[Code](#)].
- [C4] **Jaemin Cho**, Jie Lei, Hao Tan, and Mohit Bansal. “Unifying Vision-and-Language Tasks via Text Generation”. *ICML*. 2021. [[Paper](#)] [[Code](#)].
- [C3] **Jaemin Cho**, Jiasen Lu, Dustin Schwenk, Hannaneh Hajishirzi, and Aniruddha Kembhavi. “X-LXMERT: Paint, Caption and Answer Questions with Multi-Modal Transformers”. *EMNLP (long)*. 2020. [[Project](#)] [[Paper](#)] [[Demo](#)] [[Code](#)] [[Video](#)] [[Press](#)].
- [C2] **Jaemin Cho**, Minjoon Seo, and Hannaneh Hajishirzi. “Mixture Content Selection for Diverse Sequence Generation”. *EMNLP (long)*. 2019. [[Paper](#)] [[Code](#)].
- [C1] Yookoon Park, **Jaemin Cho**, and Gunhee Kim. “A Hierarchical Latent Structure for Variational Conversation Modeling”. *NAACL (long, oral)*. 2018. [[Paper](#)] [[Code](#)] [[Video](#)].

Workshop Papers

- [W1] **Jaemin Cho**, Linjie Li, Zhengyuan Yang, Zhe Gan, Lijuan Wang, and Mohit Bansal. “Diagnostic Benchmark and Iterative Inpainting for Layout-Guided Image Generation”. *The First Workshop on the Evaluation of Generative Foundation Models*. 2024. [[Project](#)] [[Paper](#)] [[Code](#)].

Preprints

- [P6] Han Lin*, **Jaemin Cho***, Abhay Zala, and Mohit Bansal. “Ctrl-Adapter: An Efficient and Versatile Framework for Adapting Diverse Controls to Any Diffusion Model”. 2024. [[Project](#)] [[Paper](#)] [[Code](#)].
- [P5] Abhay Zala*, **Jaemin Cho***, Han Lin, Jaehong Yoon, and Mohit Bansal. “EnvGen: Generating and Adapting Environments via LLMs for Training Embodied Agents”. 2024. [[Project](#)] [[Paper](#)] [[Code](#)].
- [P4] Jialu Li*, **Jaemin Cho***, Yi-Lin Sung, Jaehong Yoon, and Mohit Bansal. “SELMA: Learning and Merging Skill-Specific Text-to-Image Experts with Auto-Generated Data”. 2024. [[Project](#)] [[Paper](#)] [[Code](#)].
- [P3] David Wan, **Jaemin Cho**, Elias Stengel-Eskin, and Mohit Bansal. “Contrastive Region Guidance: Improving Grounding in Vision-Language Models without Training”. 2024. [[Project](#)] [[Paper](#)] [[Code](#)].
- [P2] Abhay Zala, Han Lin, **Jaemin Cho**, and Mohit Bansal. “DiagrammerGPT: Generating Open-Domain, Open-Platform Diagrams via LLM Planning”. 2023. [[Project](#)] [[Paper](#)] [[Code](#)].
- [P1] Han Lin, Abhay Zala, **Jaemin Cho**, and Mohit Bansal. “VideoDirectorGPT: Consistent Multi-Scene Video Generation via LLM-Guided Planning”. 2023. [[Project](#)] [[Paper](#)] [[Code](#)].

Program Committees

Reviewer

- ACL, EMNLP, NAACL, ARR, NeurIPS, COLING, CVPR, WACV

Administrator

- PyTorch KR  (2017 - Present): One of the biggest Korean ML Communities, with 10K+ members [[Link](#)]

Organizer

- T4V: Transformers for Vision (CVPR Workshop 2023/2024) [[Link](#)]
- PyTorch KR Tutorial Competition (2018) [[Link](#)], PyTorch KR Paper Day (2018) [[Link](#)], PyTorch KR Meetup (2017) [[Link](#)]

Honors

- 2023 **Bloomberg Data Science Ph.D. Fellowship.**, [[Link](#)]
- 2018 **Kakao Travel Grant (USD \$4000)**, ICML 2018
- 2017 **Semi-finalist (Top 10)**, Amazon Alexa Prize 2017

Talks

- 2024 **Korea U.**, Language Model as your Multimodal Generator, Evaluator, and Trainer
- 2023 **UNC, Korea U.**, Hierarchical Video-Moment Retrieval and Step-Captioning
- 2022 **Naver, KAIST, Postech, Korea U., Chung-Ang U.**, Vision-and-Language: Pretraining, Transfer Learning, and Evaluation
- 2021 **Kakao Brain**, Unifying Vision-and-Language Tasks via Text Generation
- 2020 **UNC**, X-LXMERT: Paint, Caption and Answer Questions with Multi-Modal Transformers
- 2020 **SNU, UNIST**, Mixture Content Selection for Diverse Sequence Generation
- 2018 **Kakao**, Generative Models & Variational Inference
- 2017 **TensorFlow KR**, Developing Korean Chatbot 101 [[Link](#)]

References

Mohit Bansal

FULL PROFESSOR [[HOMEPAGE](#)]

UNC

mbansal@cs.unc.edu

Jason Baldridge

RESEARCH SCIENTIST [[HOMEPAGE](#)]

Google

jasonbaldridge@google.com

Zhe Gan

PRINCIPAL RESEARCHER [[HOMEPAGE](#)]

Apple
zhe.gan@apple.com

Trung Bui

RESEARCH MANAGER [[HOMEPAGE](#)]

Adobe
bui@adobe.com

Aniruddha Kembhavi

RESEARCH MANAGER [[HOMEPAGE](#)]

AI2
anik@allenai.org

Minjoon Seo

ASSISTANT PROFESSOR [[HOMEPAGE](#)]

KAIST
minjoon@kaist.ac.kr

Hannaneh Hajishirzi

ASSOCIATE PROFESSOR [[HOMEPAGE](#)]

UW
hannaneh@u.washington.edu

Gunhee Kim

ASSOCIATE PROFESSOR [[HOMEPAGE](#)]

SNU
gunhee@snu.ac.kr