Haoxuan You

Columbia University, 535 W 113th St., NYC, NY, 10025 | Website | Last Updated in Aug/2023
hy2612@columbia.edu | haoxuanyou@gmail.com | Google Scholar: BhysChMAAAAJ | +1 6462263052

Education

Columbia University   New York, NY
- Doctor of Philosophy (Ph.D.) in Computer Science   *Sep. 2019 - May. 2024 (Expected)*
- Adviser: Prof. Shih-Fu Chang (Dean of Engineering School, Fellow of NAI, AAAS, ACM and IEEE)

Xidian University   Xi'an, China
- Bachelor of Engineering in Electronic Information Engineering   *Sep. 2014 - Jun. 2018*
- GPA: 3.77/4.0

Research Interests

Vision-Language Learning; Text-to-Image Generation; Multimodal Foundation Models.

Experience

AI/ML, Apple   Seattle
Research Intern   *May. 2023 - Now*
- Advisor: Liangliang Cao, Yinfei Yang, Zhe Gan
- Working on Multimodal Foundation Model.

Google Research, Google   NYC
Student Researcher   *May. 2022 - Mar. 2023*
- Advisor: Mandy Guo, Jiahui Yu, Jason Baldridge
- Project: Large-Scale Bi-directional Image-Text Generation Model. (Train w/ 1024 TPUs/job)

Azure Cognitive Services Research, Microsoft   Remote
Research Intern   *Jun. 2021 - Aug. 2021*
- Advisor: Luowei Zhou
- Project: Modality-Shared Contrastive Image-Text Pre-training.

Publication

[1] CoBIT: A Contrastive Bi-directional Image-Text Generation Model
Haoxuan You, Mandy Guo, Zhecan Wang, Kai-Wei Chang, Jason Baldridge, Jiahui Yu
Arxiv, Under Submission, 2023

[2] IdealGPT: Iteratively Decomposing Vision and Language Reasoning via Large Language Models
Haoxuan You*, Rui Sun*, Zhecan Wang*, Long Chen, Gengyu Wang, Hammad A Ayyubi, Kai-Wei Chang, Shih-Fu Chang
Arxiv, Under Submission, 2023

[3] UniFine: A Unified and Fine-grained Approach for Zero-shot Vision-Language Understanding
Rui Sun*, Zhecan Wang*, Haoxuan You*, Noel Codella, Kai-Wei Chang, Shih-Fu Chang
*Annual Meeting of the Association for Computational Linguistics (**ACL-findings**) 2023*

[4] Learning Visual Representation from Modality-Shared Contrastive Language-Image Pre-training
Haoxuan You*, Luowei Zhou*, Bin Xiao*, Noel Codella*, Yu Cheng, Ruochen Xu, Shih-Fu Chang, Lu Yuan. (*Equal Contribution)
*European Conference on Computer Vision (**ECCV**) 2022*

[5] Find Someone Who: Visual Commonsense Understanding in Human-Centric Grounding
Haoxuan You, Rui Sun, Zhecan Wang, Kai-Wei Chang, Shih-Fu Chang.
*Findings of Empirical Methods in Natural Language Processing (**EMNLP-findings**) 2022*

[6] Understanding ME? Multimodal Evaluation for Fine-grained Visual Commonsense
*Zhecan Wang, **Haoxuan You**, Yicheng He, Wenhao Li, Kai-Wei Chang, Shih-Fu Chang*
*Empirical Methods in Natural Language Processing (**EMNLP**) 2022*

[7] Rethinking network design and local geometry in point cloud: A simple residual mlp framework
*Xu Ma, Can Qin, **Haoxuan You**, Haoxi Ran, Yun Fu.*
*International Conference on Learning Representations(**ICLR**) 2022*

[8] SGEITL: Scene Graph Enhanced Image-Text Learning for Visual Commonsense Reasoning
Zhecan Wang, **Haoxuan You***, Alireza Zareian, Liunian Li, Suji Park, Yiqing Liang, Kai-Wei Chang, Shih-Fu Chang. (*Equal Contribution)*
*36th AAAI Conference on Artificial Intelligence (**AAAI**) 2022*

[9] Unsupervised Vision-and-Language Pre-training Without Parallel Images and Captions
*Liunian Harold Li, **Haoxuan You**, Zhecan Wang, Alireza Zareian, Shih-Fu Chang, Kai-Wei Chang*
*Annual Conference of the North American Chapter of the Association for Computational Linguistics (**NAACL**) 2021*

[10] Learning Visual Commonsense for Robust Scene Graph Generation
Alireza Zareian, Zhecan Wang*, **Haoxuan You***, Shih-Fu Chang. (*Equal Contribution)*
*European Conference on Computer Vision (**ECCV**) 2020*

[11] PointHop: An Explainable Machine Learning Method for Point Cloud Classification
*Min Zhang, **Haoxuan You***, Pranav Kadam, Shan Liu, C-C Jay Kuo. (*Corresponding Author)*
IEEE Transactions on Multimedia, VOL. 22, NO. 7, JULY 2020

[12] PointDAN: A Multi-Scale 3D Domain Adaption Network for Point Cloud Representation
Can Qin, **Haoxuan You***, Lichen Wang, C-C Jay Kuo, Yun Fu. (*Equal Contribution)*
*Conference on Neural Information Processing Systems (**NeurIPS**) 2019*

[13] Decoding EEG by Visual-guided Deep Neural Networks
*Zhicheng Jiao, **Haoxuan You**, Fan Yang, Xin Li, Han Zhang, Dinggang Shen.*
*28th International Joint Conference on Artificial Intelligence (**IJCAI**) 2019*

[14] Multi-modality latent interaction network for visual question answering
*Peng Gao, **Haoxuan You**, Zhanpeng Zhang, Xiaogang Wang, Hongsheng Li.*
*IEEE International Conference on Computer Vision (**ICCV**) 2019*

[15] Dynamic Fusion with Intra- and Inter-modality Attention Flow for Visual Question Answering
*Peng Gao, Hongsheng Li, **Haoxuan You**, Zhengkai Jiang, Pan Lu, Steven Hoi, Xiaogang Wang.*
*IEEE Conference on Computer Vision and Pattern Recognition (**CVPR**) 2019*

[16] PVRNet: Point-View Relation Neural Network for 3D Shape Recognition
***Haoxuan You**, Yifan Feng, Xibin Zhao, Changqing Zou, Rongrong Ji, Yue Gao.*
*33rd AAAI Conference on Artificial Intelligence (**AAAI**) 2019*

[17] Hypergraph Neural Network
*Yifan Feng, **Haoxuan You**, Rongrong Ji, Yue Gao.*
*33rd AAAI Conference on Artificial Intelligence (**AAAI**) 2019*

[18] MeshNet: Mesh Neural Network for 3D Shape Representation
*YuTong Feng, Yifan Feng, **Haoxuan You**, Xibin Zhao, Yue Gao.*
*33rd AAAI Conference on Artificial Intelligence (**AAAI**) 2019*

[19] PVNet: A Joint Convolutional Network of Point Cloud and Multi-View for 3D Shape Recognition
***Haoxuan You**, Yifan Feng, Rongrong Ji, Yue Gao.*
*ACM International Conference on Multimedia (**ACM MM**) 2018*

Computer Programming

Programming: Python, C, C++, MATLAB and others
Tools: Pytorch, Jax, Tensorflow, Keras.