Huiyu Wang

(310) 948 - 0803
https://csrhddlam.github.io
williamwanghuiyu@gmail.com

WORK EXPERIENCE	**Facebook AI Research (FAIR), Meta**, New York, NY	Feb 2022 - present
	Senior Research Scientist	
	Research Scientist	Feb 2022 - Jun 2023
EDUCATION	**Johns Hopkins University**	Aug 2017 - Jan 2022
	Ph.D. in Computer Science	
	Advisor: Alan Yuille	
	Research Interest: Computer Vision	
	University of California, Los Angeles	Sep 2015 - Dec 2016
	M.S. in Electrical Engineering	
	Shanghai Jiao Tong University	Sep 2011 - Jun 2015
	B.S. in Information Engineering	
INTERNSHIP EXPERIENCE	**Google Research**, Remote	Aug 2019 - Aug 2021
	Student Researcher	
	Mentors: Liang-Chieh Chen, Yukun Zhu	
	Google Research, Seattle, WA	May 2019 - Aug 2019
	Research Intern	
	Mentors: Yukun Zhu, Liang-Chieh Chen	
	Allen Institute for Artificial Intelligence, Seattle, WA	May 2018 - Aug 2018
	Research Intern	
	AI2 Outstanding Intern of 2018 Award	
	Mentors: Mohammad Rastegari, Aniruddha Kembhavi, Ali Farhadi	
	Johns Hopkins University, Baltimore, MD	Feb 2017 - Jun 2017
	Research Assistant	
	Advisor: Alan Yuille	
	TuSimple, San Diego, CA	Jun 2016 - Nov 2016
	Research Engineering Intern	
	Mentor: Xiaodi Hou	
	University of California, Los Angeles, Los Angeles, CA	Apr 2016 - Jun 2016
	Research Assistant	
	Advisors: Ying Nian Wu, Song-Chun Zhu	
	Shanghai Jiao Tong University, Shanghai, China	Dec 2014 - Jun 2015
	Undergraduate Researcher	
	Advisor: Li Song	
PH.D. THESIS	**Huiyu Wang**. On Modeling Long-Range Dependencies for Visual Perception.	
	Ph.D. thesis, Johns Hopkins University, Baltimore, MD, January 2022. 230 pages.	

PUBLICATIONS   Feng Cheng, Mi Luo, **Huiyu Wang**, Alex Dimakis, Lorenzo Torresani, Gedas Bertasius, and Kristen Grauman. 4Diff: 3D-Aware Diffusion Model for Third-to-First Viewpoint Translation. In *Proceedings of the 18th European Conference on Computer Vision (ECCV)*, Milano, Italy, September 29 - October 4, 2024.

Wufei Ma, Kai Li, Zhongshi Jiang, Moustafa Meshry, Qihao Liu, **Huiyu Wang**, Christian Haene, and Alan Yuille. Rethinking Video-Text Understanding: Retrieval from Counterfactually Augmented Data. In *Proceedings of the 18th European Conference on Computer Vision (ECCV)*, Milano, Italy, September 29 - October 4, 2024.

Md Mohaiminul Islam, Tushar Nagarajan, **Huiyu Wang**, Fu-Jen Chu, Kris Kitani, Gedas Bertasius, and Xitong Yang. Propose, Assess, Search: Harnessing LLMs for Goal-Oriented Planning in Instructional Videos. In *Proceedings of the 18th European Conference on Computer Vision (ECCV)*, Milano, Italy, September 29 - October 4, 2024.

K R Prajwal, Bowen Shi, Matthew Le, Apoorv Vyas, Andros Tjandra, Mahi Luthra, Baishan Guo, **Huiyu Wang**, Triantafyllos Afouras, David Kant, and Wei-Ning Hsu. MusicFlow: Cascaded Flow Matching for Text Guided Music Generation. In *Proceedings of the 41st International Conference on Machine Learning (ICML)*, Vienna, Austria, PMLR 235, July 21-27, 2024.

Kristen Grauman, Andrew Westbury, Lorenzo Torresani, Kris Kitani, Jitendra Malik, Triantafyllos Afouras, Kumar Ashutosh, Vijay Baiyya, Siddhant Bansal, Bikram Boote, Eugene Byrne, Zach Chavis, Joya Chen, Feng Cheng, Fu-Jen Chu, Sean Crane, Avijit Dasgupta, Jing Dong, Maria Escobar, Cristhian Forigua, Abrham Gebreselasie, Sanjay Haresh, Jing Huang, Md Mohaiminul Islam, Suyog Jain, Rawal Khirodkar, Devansh Kukreja, Kevin J Liang, Jia-Wei Liu, Sagnik Majumder, Yongsen Mao, Miguel Martin, Effrosyni Mavroudi, Tushar Nagarajan, Francesco Ragusa, Santhosh Kumar Ramakrishnan, Luigi Seminara, Arjun Somayazulu, Yale Song, Shan Su, Zihui Xue, Edward Zhang, Jinxu Zhang, Angela Castillo, Changan Chen, Xinzhu Fu, Ryosuke Furuta, Cristina Gonzalez, Prince Gupta, Jiabo Hu, Yifei Huang, Yiming Huang, Weslie Khoo, Anush Kumar, Robert Kuo, Sach Lakhavani, Miao Liu, Mi Luo, Zhengyi Luo, Brighid Meredith, Austin Miller, Oluwatumininu Oguntola, Xiaqing Pan, Penny Peng, Shraman Pramanick, Merey Ramazanova, Fiona Ryan, Wei Shan, Kiran Somasundaram, Chenan Song, Audrey Southerland, Masatoshi Tateno, **Huiyu Wang**, Yuchen Wang, Takuma Yagi, Mingfei Yan, Xitong Yang, Zecheng Yu, Shengxin Cindy Zha, Chen Zhao, Ziwei Zhao, Zhifan Zhu, Jeff Zhuo, Pablo Arbelaez, Gedas Bertasius, David Crandall, Dima Damen, Jakob Engel, Giovanni Maria Farinella, Antonino Furnari, Bernard Ghanem, Judy Hoffman, C. V. Jawahar, Richard Newcombe, Hyun Soo Park, James M. Rehg, Yoichi Sato, Manolis Savva, Jianbo Shi, Mike Zheng Shou, and Michael Wray. Ego-Exo4D: Understanding Skilled Human Activity from First- and Third-Person Perspectives. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 19383-19400, Seattle, WA, June 17-21, 2024. **(Oral)**

Yuhan Shen, **Huiyu Wang**, Xitong Yang, Matt Feiszli, Ehsan Elhamifar, Lorenzo Torresani, and Effrosyni Mavroudi. Learning to Segment Referred Objects from Narrated Egocentric Videos. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 14510-14520, Seattle, WA, June 17-21, 2024. **(Oral)**

Yale Song, Gene Byrne, Tushar Nagarajan, **Huiyu Wang**, Miguel Martin, and Lorenzo Torresani. Ego4D Goal-Step: Toward Hierarchical Understanding of Procedural Ac-

tivities. In Alice Oh, Tristan Naumann, Amir Globerson, Kate Saenko, Moritz Hardt, and Sergey Levine, editors, volume 36 of *Advances in Neural Information Processing Systems: Annual Conference on Neural Information Processing Systems (NeurIPS)*, pages 38863–38886, New Orleans, LA, Curran Associates, Inc., December 10-16, 2023. **(Spotlight)**

Triantafyllos Afouras, Effrosyni Mavroudi, Tushar Nagarajan, **Huiyu Wang**, and Lorenzo Torresani. HT-Step: Aligning Instructional Articles with How-To Videos. In Alice Oh, Tristan Naumann, Amir Globerson, Kate Saenko, Moritz Hardt, and Sergey Levine, editors, volume 36 of *Advances in Neural Information Processing Systems: Annual Conference on Neural Information Processing Systems (NeurIPS)*, pages 50310-50326, New Orleans, LA, Curran Associates, Inc., December 10-16, 2023.

Huiyu Wang, Mitesh Kumar Singh, and Lorenzo Torresani. Ego-Only: Egocentric Action Detection without Exocentric Transferring. In *Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV)*, pages 5227–5238, Paris, France, October 1-6, 2023.

Chen Wei, Karttikeya Mangalam, Po-Yao Huang, Yanghao Li, Haoqi Fan, Hu Xu, **Huiyu Wang**, Cihang Xie, Alan Yuille, and Christoph Feichtenhofer. Diffusion Models as Masked Autoencoders. In *Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV)*, pages 16238–16248, Paris, France, October 1-6, 2023.

Yuanze Lin, Chen Wei, **Huiyu Wang**, Alan Yuille, and Cihang Xie. SMAUG: Sparse Masked Autoencoder for Efficient Video-Language Pre-training. In *Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV)*, pages 2459–2469, Paris, France, October 1-6, 2023.

Boxiang Yun, Baiying Lei, Jieneng Chen, **Huiyu Wang**, Song Qiu, Wei Shen, Qingli Li, and Yan Wang. SpecTr: Spectral Transformer for Microscopic Hyperspectral Pathology Image Segmentation. *IEEE Transactions on Circuits and Systems for Video Technology*, 2023.

Yutong Bai, Zeyu Wang, Junfei Xiao, Chen Wei, **Huiyu Wang**, Alan Yuille, Yuyin Zhou, and Cihang Xie. Masked Autoencoders Enable Efficient Knowledge Distillers. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 24256–24265, Vancouver, BC, Canada, June 17-24, 2023.

Qihang Yu, **Huiyu Wang**, Siyuan Qiao, Maxwell Collins, Yukun Zhu, Hartwig Adam, Alan Yuille, and Liang-Chieh Chen. k-means Mask Transformer. In Shai Avidan, Gabriel J. Brostow, Moustapha Cissé, Giovanni Maria Farinella, and Tal Hassner, editors, *Proceedings of the 17th European Conference on Computer Vision (ECCV)*, volume 13689 of *Lecture Notes in Computer Science*, pages 288-307, Tel Aviv, Israel, Springer, Cham, October 23-27, 2022.

Feng Wang, **Huiyu Wang**, Chen Wei, Alan Yuille, and Wei Shen. CP2: Copy-Paste Contrastive Pretraining for Semantic Segmentation. In Shai Avidan, Gabriel J. Brostow, Moustapha Cissé, Giovanni Maria Farinella, and Tal Hassner, editors, *Proceedings of the 17th European Conference on Computer Vision (ECCV)*, volume 13690 of *Lecture Notes in Computer Science*, pages 499-515, Tel Aviv, Israel, Springer, Cham, October 23-27, 2022.

Xianhang Li, **Huiyu Wang**, Chen Wei, Jieru Mei, Alan Yuille, Yuyin Zhou, and Cihang Xie. In Defense of Image Pre-Training for Spatiotemporal Recognition. In

Shai Avidan, Gabriel J. Brostow, Moustapha Cissé, Giovanni Maria Farinella, and Tal Hassner, editors, *Proceedings of the 17th European Conference on Computer Vision (ECCV)*, volume 13685 of *Lecture Notes in Computer Science*, pages 675-691, Tel Aviv, Israel, Springer, Cham, October 23-27, 2022.

Qihang Yu, **Huiyu Wang**, Dahun Kim, Siyuan Qiao, Maxwell Collins, Yukun Zhu, Hartwig Adam, Alan Yuille, and Liang-Chieh Chen. CMT-DeepLab: Clustering Mask Transformers for Panoptic Segmentation. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 2560-2570, New Orleans, LA, June 18-24, 2022. **(Oral)**

Dahun Kim, Jun Xie, **Huiyu Wang**, Siyuan Qiao, Qihang Yu, Hong-Seok Kim, Hartwig Adam, In So Kweon, and Liang-Chieh Chen. TubeFormer-DeepLab: Video Mask Transformer. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 13914-13924, New Orleans, LA, June 18-24, 2022.

Sucheng Ren, **Huiyu Wang**, Zhengqi Gao, Shengfeng He, Alan Yuille, Yuyin Zhou, and Cihang Xie. A Simple Data Mixing Prior for Improving Self-Supervised Learning. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 14595-14604, New Orleans, LA, June 18-24, 2022.

Jinghao Zhou, Chen Wei, **Huiyu Wang**, Wei Shen, Cihang Xie, Alan Yuille, and Tao Kong. iBOT: Image BERT Pre-Training with Online Tokenizer. In *Proceedings of the Tenth International Conference on Learning Representations (ICLR)*, Virtual, OpenReview.net, April 25-29, 2022.

Huaijin Pi, **Huiyu Wang**, Yingwei Li, Zizhang Li, and Alan Yuille. Searching for TrioNet: Combining Convolution with Local and Global Self-Attention. In *Proceedings of the 32nd British Machine Vision Conference (BMVC)*, page 141, Virtual, BMVA Press, November 22-25, 2021.

Huiyu Wang, Yukun Zhu, Hartwig Adam, Alan Yuille, and Liang-Chieh Chen. MaX-DeepLab: End-to-End Panoptic Segmentation with Mask Transformers. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 5463-5474, Virtual, Nashville, TN, June 19-25, 2021.

Chen Wei, **Huiyu Wang**, Wei Shen, and Alan Yuille. CO2: Consistent Contrast for Unsupervised Visual Representation Learning. In *Proceedings of the 9th International Conference on Learning Representations (ICLR)*, Virtual, Vienna, Austria, OpenReview.net, May 3-7, 2021.

Huiyu Wang, Yukun Zhu, Bradley Green, Hartwig Adam, Alan Yuille, and Liang-Chieh Chen. Axial-DeepLab: Stand-Alone Axial-Attention for Panoptic Segmentation. In Andrea Vedaldi, Horst Bischof, Thomas Brox, and Jan-Michael Frahm, editors, *Proceedings of the 16th European Conference on Computer Vision (ECCV)*, volume 12349 of *Lecture Notes in Computer Science*, pages 108-126, Glasgow, UK, Springer, Cham, August 23-28, 2020. **(Spotlight)**

Adam Kortylewski, Qing Liu, **Huiyu Wang**, Zhishuai Zhang, and Alan Yuille. Combining Compositional Models and Deep Networks For Robust Object Classification under Occlusion. In *Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision (WACV)*, pages 1333-1341, Snowmass Village, CO, March 1-5, 2020. **(Spotlight)**

Qing Liu, Lingxi Xie, **Huiyu Wang**, and Alan Yuille. Semantic-Aware Knowledge Preservation for Zero-Shot Sketch-Based Image Retrieval. In *Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV)*, pages 3662-3671, Seoul, South Korea, October 27 - November 2, 2019.

Adam Kortylewski, Qing Liu, **Huiyu Wang**, Zhishuai Zhang, and Alan Yuille. Localizing Occluders with Compositional Convolutional Networks. In *Proceedings of the IEEE/CVF International Conference on Computer Vision Workshop (ICCVW)*, pages 2029-2032, Seoul, South Korea, October 27-28, 2019.

Huiyu Wang, Aniruddha Kembhavi, Ali Farhadi, Alan Yuille, and Mohammad Rastegari. ELASTIC: Improving CNNs with Dynamic Scaling Policies. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 2258-2267, Long Beach, CA, June 16-20, 2019. **(Oral)**

Zhixuan Wei, Weidong Chen, Jingchuan Wang, **Huiyu Wang**, and Kang Li. Semantic Mapping for Safe and Comfortable Navigation of a Brain-Controlled Wheelchair. In Jangmyung Lee, Min Cheol Lee, Honghai Liu, and Jee-Hwan Ryu, editors, *Proceedings of the 6th International Conference on Intelligent Robotics and Applications (ICIRA)*, volume 8102 of *Lecture Notes in Computer Science*, pages 307-317, Busan, South Korea, Springer, Berlin, Heidelberg, September 25-28, 2013.

TECHNICAL REPORTS	Mark Weber[*], **Huiyu Wang**[*], Siyuan Qiao[*], Jun Xie, Maxwell D. Collins, Yukun Zhu, Liangzhe Yuan, Dahun Kim, Qihang Yu, Daniel Cremers, Laura Leal-Taixe, Alan Yuille, Florian Schroff, Hartwig Adam, and Liang-Chieh Chen. DeepLab2: A TensorFlow Library for Deep Labeling. *Computing Research Repository*, arXiv:2106.09748, June 2021.

Liang-Chieh Chen, **Huiyu Wang**, and Siyuan Qiao. Scaling Wide Residual Networks for Panoptic Segmentation. *Computing Research Repository*, arXiv:2011.11675, November 2020.

Siyuan Qiao, **Huiyu Wang**, Chenxi Liu, Wei Shen, and Alan Yuille. Micro-batch training with batch-channel normalization and weight standardization. *Computing Research Repository*, arXiv:1903.10520, August 2020.

Siyuan Qiao, **Huiyu Wang**, Chenxi Liu, Wei Shen, and Alan Yuille. Rethinking Normalization and Elimination Singularity in Neural Networks. *Computing Research Repository*, arXiv:1911.09738, November 2019.

Siyuan Qiao, **Huiyu Wang**, Chenxi Liu, Wei Shen, and Alan Yuille. Weight Standardization. *Computing Research Repository*, arXiv:1903.10520, March 2019.

FELLOWSHIPS AND AWARDS	Top Reviewer, NeurIPS	2022
	Outstanding Reviewer, CVPR	2021
	Outstanding Reviewer, ECCV	2020
	AI2 Outstanding Intern	2018
	The SCSK® Scholarship	2014
	2nd prize, National Undergraduate Electronic Design Contest	2013
	1st prize, TI® Cup Electronic Design Contest of SJTU	2013
	Academic Excellence Scholarship, 2nd class, SJTU	2012, 2013, 2014

SKILLS

Programming Languages:	Python, C++, Java, Matlab, C#, LaTeX
Deep Learning Tools:	PyTorch, TensorFlow, MXNet, Caffe

July 2024