Zhecan (James) Wang
王哲灿

Email: olinzhecanwang@gmail.com
Semantic Scholar: Zhecan-Wang/2513111
Google Scholar ID: uqHPnmgAAAAJ
Personal Website: zhecanwang.com
GitHub: github.com/ZhecanJamesWang
Patent: patents.justia.com/inventor/zhecan-wang
Linkedin: https://www.linkedin.com/in/jameszhecanwang/

AI, Multimodal Learning, Commonsense Reasoning, Vision-Language Human-in-the-loop, Human-Centered AI, Human-AI Interaction

EDUCATION

Columbia University New York, NY

Ph.D. candidate in Computer Science *'19 - '24*

Co-supervised by **Prof.** Shih-Fu Chang (Dean of Columbia Engineering School, Fellow of ACM, National Academy of Engineering) and **Prof.** Kai-Wei Chang (UCLA, Amazon Scholar)

Ph.D. Research funded by DARPA Machine Common Sense (MCS) Program

Franklin W. Olin College of Engineering Boston, MA

Bachelor of Science in Engineering, with a Concentration in Computing *'14 - '18*

National University of Singapore (NUS) Singapore

Exchange Student, UROP Research *'16 - '17*

Co-supervised by **Prof.** Jiashi Feng and **Prof.** Ashraf Kassim (Former Provost of NUS)

Collaborated with **Prof.** Shuicheng Yan (Fellow of Singapore's Academy of Engineering, ACM, Former VP of 360)

Cross-Registration

Boston, MA

- **Wellesley College** *'14 - '18*
- **Massachusetts Institute of Technology (MIT)**
- **Babson College**

HONOR

- **1st place** in ICCV Microsoft Global Challenge of Recognizing One Million Celebrities in the Real World *'17*
  - MS-Celeb-1M, the largest data set for large-scale human face classification;
  - 1st place in all sub-categories including full-shot and low-shot problems.
- Recipient of AAAI-22 Student Scholarship Program. *'22*
- Recipient of 4-year Olin College Merit Scholarship - **$100,000**. *'14 - '18*
- **1st** place at the 2nd Facial Landmark Localization Competition *'17*
  - Hosted by iBUG (Intelligent Behaviour Understanding Group), Imperial College London.
- Invited to meet the **Prime Minister** of Denmark and participate in the cultural exhibition. *'10*

LEADERSHIP

- Lead Department Representative at the Engineering Graduate Student Council (EGSC). '19 - '21
- Research Committee Member In Transcription Club '20 - '22
  (Community for life science scientists and entrepreneurs)
- Co-founder of HackIT, an online non-profit engineering education platform '17 - '18

PROFESSIONAL ACTIVITY

- Reviewer of ICCV 2019, NeurIPS 2019, AAAI 2019, CVPR 2020, ECCV 2020, AAAI 2020, NeurIPS 2020, WACV 2020, ACCV 2020, AAAI 2021, CVPR 2021, ICCV 2021, ARR 2021, NeurIPS 2022, AAAI 2022, WACV 2022, EMNLP 2022, CVPR 2022, MM 2022, ARR 2022, ICLR 2023, AAAI 2023, MM 2023, NeurIPS 2023, AAAI 2023, WACV 2023, EMNLP 2023, ARR 2023, AAAI 2024, WACV 2024 and ICLR 2024.
- IEEE (Institute of Electrical and Electronics Engineers) Membership '18 - Present

INDUSTRY

❖ **Google Deepmind (Google Brain)**, supervised by Dr. Golnaz Ghiasi & Dr. Adam Yu) Mountain View, CA
- Graduate Research Intern *Jun. '22 - Sep. '22*
- Part-time Student Researcher *Sep. '22 - Aug. '23*
- Research on vision language understanding, including arithmetic reasoning, causal relationships, etc.

❖ **Microsoft Research** (supervised by Dr. Noel Codella, Principle Researcher) Redmond, WA
- Graduate Research Intern *Jun. '21- Sep. '21*
- Part-time Researcher *Sep. '21 - Mar. '22*
- Research on Multimodal AI includes multimedia like vision language;
- Utilizing pre-trained vision language models for downstream tasks.

❖ **Xpeng Motors** (supervised by Dr. Yandong Guo, current VP of OPPO) Mountain View, CA
- A.I. Research Engineer *Jul. '18- Aug. '19*
- Research on human-machine interaction;
- The main contributor to the Smart Cabin project included analyzing driver's attention, fatigue, heart rate, and behavior;
- The only undergraduate employee in the research center and directly works with the Chief Scientist, **Dr.** Yandong Guo, former MSR researcher and current VP of OPPO.

❖ **Panasonic R&D Center,** Learning & Vision Lab Singapore
- ML Research Intern *Jun. '17 - Aug. '17*
- The only undergraduate in the 4-person core development team, MS-Celeb-1M (1st place);
- Led the undergraduate intern team for data processing under the supervision of **Prof.** Jian Zhao and **Prof.** Jiashi Feng;
- Publications at ICCV 2017, BMVC 2017.

❖ **IVANI** (Smart home device startup) Saint Louis, IL

- Data Science and Back-End Dev Intern *May '16 - Aug. '16*
- Non-invasive human body localization and activity detection;
- Developed automated sensor data collection pipeline, data analysis, visualization, and algorithms;
- Worked directly with senior engineers and demoed to the board and clients;
- My work is under 7 Patents and surpassed the company's main algorithm.

ACADEMIC

- **Columbia University, DVMM Lab**, Graduate Research Assistant New York, NY
  (Co-supervised by **Prof.** Shih-Fu Chang and Prof. Kai-Wei Chang) *Sep. '19 - May '24*
  - Learning with Multimedia (image, video, text, etc.);
  - **My Work:** Multimodal ML, commonsense reasoning, and human/brain machine interaction.

- **National University of Singapore, LV Lab**, Research Assistant Singapore
  (Co-supervised by **Prof.** Jiashi Feng and **Prof.** Ashraf Kassim) *Jan. '17 - May '17*
  - One of the top AI labs in Asia. Face-related Research;
  - **My Work:** Occluded Face Detection, Neural Style Transfer, and Case Sensitive Metric Learning;

- **Massachusetts Institute of Technology (MIT), Media Lab**, Research Assistant Boston, MA
  (Supervised by **Prof.** Deb Roy) *Jun. '16 - Dec. '16*
  - Modeling Dialogue Structure in SMS Messages
  - Developed a computational tool to automatically model dialogue structure in SMS messages and help users better understand large SMS conversations;
  - **My Work:** Data collecting, database management, model implementation, embedding evaluation.

- **Olin College of Engineering, OCCaM lab**, Research Assistant Needham, MA
  (Supervised by **Prof.** Paul Ruvolo) *Jun. '15 - Aug. '15*
  - Eye Helper: Non-invasive, assistive technology to help people with visual impairment shop for groceries;
  - **My Work:** Created the infrastructure, including data processing pipelines on ROS, conducted sensor fusion for local navigation, and created a wearable interface with motion sensors, socket connection, voice recognition.
  - Lifelong Turtlebot: Taught a Turtlebot to continuously learn real-life objects with RGB-D cameras along with lidar, laser, etc.
  - **My Work:** Designed and implemented infrastructure for data collection and models' training. Implemented algorithms for deployment.

PUBLICATION

- MuSiC: Machine-Human-In-the-Loop System for Visual Commonsense Challenges
  Zhecan Wang, Golnaz Ghiasi, Adams Yu, Thang Luong, Kai-Wei Chang, Shih-fu Chang, Quoc V. Le

CVPR 2024 Submission, The IEEE/CVF Conference on Computer Vision and Pattern Recognition 2024.

- Adversarial Complex Hallucination Challenges in Visual Question Answering
  Zhecan Wang, Golnaz Ghiasi, Adams Yu, Thang Luong, Kai-Wei Chang, Shih-fu Chang, Quoc V. Le
  CVPR 2024 Submission, The IEEE/CVF Conference on Computer Vision and Pattern Recognition 2024.
- MAD: Multimodal Adaptive Distillation for Leveraging Unimodal Encoders for Vision-Language Tasks
  Zhecan Wang*, Noel Codella*, Yen-Chun Chen, Luowei Zhou, Xiyang Dai, Bin Xiao, Jianwei Yang, Haoxuan You, Kai-Wei Chang, Shih-fu Chang, Lu Yuan
  IJCAI-PRICAI '24 Submission, International Joint Conference on Artificial Intelligence and Pacific Rim International Conference on Artificial Intelligence 2024.
  Former CLIP-TD, SOTA performance on VCR, VQA, and SNLI-VE benchmarks.
- CoBIT: A Contrastive Bi-directional Image-Text Generation Modes
  Haoxuan You, Mandy Guo, **Zhecan Wang**, Kai-Wei Chang, Jason Baldridge, Jiahui Yu
  ICLR '24 Submission, The Twelfth International Conference on Learning Representations 2024.
- Learning Knowledge-Aware Multimodal Representation for Visual Commonsense Reasoning
  Zhecan Wang, Haoxuan You, Alireza Zareian, Kai-Wei Chang, Shih-Fu Chang
  IJCAI-PRICAI '24 Submission, International Joint Conference on Artificial Intelligence and Pacific Rim International Conference on Artificial Intelligence 2024.
- Dataset Bias Mitigation in Multiple-Choice Visual Question Answering and Beyond
  Zhecan Wang, Long Chen, Haoxuan You, Keyang Xu, Yicheng He, Wenhao Li, Noel Codella, Kai-Wei Chang, Shih-Fu Chang
  EMNLP '23 Findings, Proceedings of the 2023 Conference on Empirical Methods in Natural Language Processing
- IdealGPT: Iteratively Decomposing Vision and Language Reasoning via Large Language Models
  Haoxuan You*, **Zhecan Wang***, Rui Sun*, Long Chen, Gengyu Wang, Hammad A. Ayyubi, Kai-Wei Chang, Shih-Fu Chang
  EMNLP '23 Findings, Proceedings of the 2023 Conference on Empirical Methods in Natural Language Processing
- UniFine: A Unified and Fine-grained Approach for Zero-shot Vision-Language Understanding
  Zhecan Wang*, Rui Sun*, Haoxuan You, Noel Codella, Kai-Wei Chang, Shih-Fu Chang
  ACL '23 Findings, Proceedings of the 61th Annual Meeting of the Association for Computational Linguistics
- Understanding ME? Multimodal Evaluation for Fine-grained Visual Commonsense
  Zhecan Wang, Haoxuan You, Yicheng He, Wenhao Li, Kai-Wei Chang, and Shih-Fu Chang
  EMNLP '22, Proceedings of the 2022 Conference on Empirical Methods in Natural Language Processing
- Find Someone Who: Visual Commonsense Understanding in Human-Centric Grounding
  Haoxuan You*, Rui Sun*, **Zhecan Wang**, Kai-Wei Chang, Shih-Fu Chang. *EMNLP '22 Findings, Proceedings of the 2022 Conference on Empirical Methods in Natural Language Processing Findings*
- SGEITL: Scene Graph Enhanced Image-Text Learning for Visual Commonsense Reasoning
  Zhecan Wang*, Haoxuan You*, Liunian Harold Li, Alireza Zareian, Suji Park, Yiqing Liang, Kai-Wei Chang, Shih-Fu Chang
  AAAI '22, The 36th AAAI Conference on Artificial Intelligence 2022.
- Graph-MLP: Node Classification without Message Passing in Graph
  Yang Hu, Haoxuan You, **Zhecan Wang**, Zhicheng Wang, Erjin Zhou, Yue Gao
  2021 arXiv preprint: 2106.04051
- Bridging the Gap between Recognition-level Pre-training and Cognition-level Vision-language Tasks
  Yue Wan*, Yueen Ma*, Haoxuan You, **Zhecan Wang**, and Shih-Fu Chang.
  ACL-CSRR '22, Proceedings of the 60th Annual Meeting of the Association for Computational Linguistics

- Weakly-supervised VisualBERT: Pre-training without Parallel Images and Captions
  Liunian Harold Li, Haoxuan You*, **Zhecan Wang***, Alireza Zareian, Shih-Fu Chang, Kai-Wei Chang
  NAACL ' 21, 2021 Annual Conference of the North American Chapter of the Association for Computational Linguistics.
- Learning Visual Commonsense for Robust Scene Graph Generation
  Alireza Zareian*, **Zhecan Wang***, Haoxuan You*, Shih-Fu Chang
  ECCV '20, 2020 European Conference on Computer Vision (Oral).
- Learning to Detect Head Movement in Unconstrained Remote Gaze Estimation in the Wild
  Zhecan Wang, Jian Zhao, Cheng Lu, Fan Yang, Han Huang, lianji li, Yandong Guo
  WACV '20, 2020 IEEE/CVF Winter Conference on Applications of Computer Vision (Oral).
- Conditional Dual-Agent GANs for Photorealistic and Annotation Preserving Image Synthesis
  Zhecan Wang, Jian Zhao, Yu Cheng, Shengtao Xiao, Jianshu Li, Fang Zhao, Jiashi Feng, Ashraf Kassim
  BMVCW ' 17, 28th British Machine Vision Conference Face HUB WorkShop, 2017 (Oral), Travel grant.
- Know You at One Glance: A Compact Vector Representation for Low-Shot Learning
  Yu Cheng*, Jian Zhao*, **Zhecan Wang,** Yan Xu, Karlekar Jayashree, Shengmei Shen, Jiashi Feng
  ICCVW '17, 2017 IEEE International Conference on Computer Vision Workshops (Oral).
- High Performance Large Scale Face Recognition with Multi-Cognition Softmax and Feature Retrieval
  Yan Xu*, Yu Cheng*, Jian Zhao, **Zhecan Wang**, Lin Xiong, J. Karlekar, Hajime Tamura, Tomoyuki Kagaya, Sugiri Pranata, Shengmei Shen, Jiashi Feng, Junliang Xing
  ICCVW '17, 2017 IEEE International Conference on Computer Vision Workshops (Oral).
- 3D-assisted Coarse-to-fine Extreme-pose Facial Landmark Detection
  Shengtao Xiao, Jianshu Li, Yunpeng Chen, **Zhecan Wang**, Jiashi Feng, Shuicheng Yan, A. Kassim
  CVPRW '17, 2017 IEEE Conference on Computer Vision and Pattern Recognition Workshops (Poster).
- Dual-Agent GANs for Photorealistic and Identity Preserving Profile Face Synthesis
  Jian Zhao, Lin Xiong, J. Karlekar, Jianshu Li, F. Zhao, **Zhecan Wang**, Sugiri Pranata, Shengmei Shen, Shuicheng Yan, Jiashi Feng
  NIPS '17, Proceedings of the 31st International Conference on Neural Information Processing Systems
- Self-Supervised Neural Aggregation Networks for Human Parsing
  Jian Zhao, Jianshu Li, Xuecheng Nie, Fang Zhao, Yunpeng Chen, **Zhecan Wang**, Jiashi Feng, Shuicheng Yan
  CVPRW '17, 2017 IEEE Conference on Computer Vision and Pattern Recognition Workshops (Oral).
- Autonomous Vehicles for Remote Sample Collection in Difficult Conditions: Enabling remote sample collection by marine biologists
  Andrew Bennett, Victoria Preston, Jay Woo, Shivali Chandra, Devynn Diggins, Riley Chapman, **Zhecan Wang**, Matthew Rush, Liani Lye, Mindy Tieu, Silas Hughes, Iain Kerr, Adela Wee
  TePRA '15, 2015 IEEE International Conference on Technologies for Practical Robot Applications
- Autonomous Vehicles for Remote Sample Collection: Enabling Marine Research
  Andrew Bennett, David Barrett, Victoria Preston, Jay Woo, Shivali Chandra, Devynn Diggins, Riley Chapman, Adela Wee, **Zhecan Wang**, Matthew Rush, Iain Kerr
  OCEANS '15, 2015 IEEE International Conference on OCEANS – Genova.

PATENT

7 US Utility Patents about Indoor Human Body Analysis with Wireless Sensors and Artificial Intelligence

1 CN Patent about Human Attention Detection via Artificial Intelligence

- Detecting location within a network. Patent No. US11423968. August 23, 2022.
- Detecting location within a network. Patent No. US10964371. March 30, 2021.
- Detecting location within a network. Patent No. US10665284. May 26, 2020.
- Detecting location within a network. Patent No. US10455357. October 22, 2019.
- Detecting location within a network. Patent No. US10325641. June 18, 2019.
- Detecting location within a network. Patent No. US10142785. November 27, 2018.
- Detecting location within a network. Patent No. US10064014. August 28, 2018.
- Human eye sight line direction determination method and system of human eye sight line detection system. Patent No. CN110458104B. December 7. 2021.

TEACHING

Head Teaching Assistant, Fall 2022, COMS W4995: Neural Network and Deep Learning
Teaching Assistant, Spring 2022, COMS 6998: Advanced Topics in Spoken Language Processing

PERSONAL PROJECT

- Improving Human-Robot Collaboration through Wearable Technology *Sep. '17 - May '18*
  - Sponsored by **Rockwell Automation** (**top 1** robotic arm company) and protected under **NDA**;
  - Using wearable technology to harness emotional intelligence between humans and industrial robots to increase interaction while maintaining safety standards.

- Neato Robotics Street View *Sep. '16 - Dec. '16*
  - Self-made "Google street 3D view" on campus with a vacuum robot in Python and JavaScript;
  - Implemented SLAM (simultaneous localization and mapping) algorithm with our version of the particle filter localization; method and used that algorithm to navigate robots around for data collecting;
  - Created a panorama view of multiple locations by image-stitching in OpenCV; developed a web interface for users.

- Assisting people with disabilities with image and word classification via brain wave *Jan. '22 - Mar. '22*
  - Collaborated with Microsoft Research researchers;
  - Using machine learning algorithm to analyze EEG brain wave signals.

SKILL

Programming: Machine Learning, Deep Learning, Robotics Controlling, Sensor Fusion, User Interface Design, Model Deployment, Dataset Management, Software Design, Software Development, App Development, Website Development, Microcontroller and Sensor, Mechanical Design, Control System Development, Prototype Design, Process Control, Simulation Software, Embedded System, etc.

REFERENCE

References are listed by priority.

Name: Shih-fu Chang
Relationship: Main-supervisor
Position: Professor, Dean of Columbia Engineering
Employer: Columbia University
Email: sc250@columbia.edu
Dossier Email: send.Engineering.1047E47D9F@interfoliodossier.com

Name: Kai-Wei Chang
Relationship: Co-supervisor
Position: Associate Professor
Employer: University of California, Los Angeles (UCLA)
Email: kwchang@cs.ucla.edu
Dossier Email: send.Engineering.DE4058A6A1@interfoliodossier.com

Name: Noel Codella
Relationship: Internship Mentor
Position: Principal Researcher
Employer: Microsoft Research (MSR)
Email: ncodella@microsoft.com
Dossier Email: send.Engineering.55151F0851@interfoliodossier.com

Name: Ashraf Kassim
Relationship: Provost
Position: Professor & Associate Provost (Adjunct Professor and former Vice Provost at NUS)
Employer: Singapore University of Technology and Design (SUTD)
Email: ashraf@sutd.edu.sg, ashraf@nus.edu.sg
Dossier Email: send.Kassim.1D777D9D66@interfoliodossier.com

Name: Long Chen
Relationship: Research Collaborator
Position: Assistant Professor
Employer: The Hong Kong University of Science and Technology (HKUST)
Email: longchen@ust.hk
Dossier Email: send.Engineering.0FE9F1C98F@interfoliodossier.com