
EDUCATION

- **Carnegie Mellon University** Pittsburgh, PA
Master of Science in Computer Vision 08/13/2018 – 12/18/2019
- **Georgia Institute of Technology** Atlanta, GA
B.S. in Computer Science, B.S. in Industrial and System Engineering; 08/19/2013 – 05/04/2018
Honors: Summa Cum Laude

INDUSTRY EXPERIENCE

- **Amazon Web Services, Inc.** Seattle, WA
Applied Scientist 03/09/2020 - Present
 - Amazon Artificial General Intelligence
 - Launched Amazon Titan (2023) and Amazon Nova (2024) suites of foundation models, including Amazon’s large language models (LLMs), image generation models, and video generation models.
 - Responsible for R&D to improve the performance, safety, and transparency of generative AI models.
 - Amazon AutoGluon
 - Core contributor to AutoGluon, an AutoML library that automates machine learning for applications with only a few lines of code.
 - Integrating AutoGluon to power AWS SageMaker AutoPilot which allows users to build and deploy models with no ML knowledge and no code
 - github repository has 8000+ stars and 900+ forks
 - Amazon Rekognition
 - Designed and implemented novel AI/ML algorithms for face recognition technology
 - Published [4] [3] [2] in ICCV 2021 and ECCV 2022
- **Carnegie Mellon University** Pittsburgh, PA
Capstone Project with Prof. Kris Kitani 08/20/2018 - 03/06/2020
 - Explored using Graph Neural Networks for Multi-Object Tracking (MOT)
 - Proposed the first GNN method for simultaneous detection and association for multi-object tracking
 - Achieved State-of-the-art performance on MOTChallenge
 - Paper [6] published to International Conference on Robotics and Automation (ICRA) 2021
 - Paper [9] published to Computer Vision and Pattern Recognition (CVPR) 2020
- **Carnegie Mellon University** Pittsburgh, PA
Research Assistant with Prof. Louis-Philippe Morency 08/12/2019 - 03/06/2020
 - Studied multimodal machine learning for modeling human communication dynamics [5]
 - Paper [5] published to NAACL-HLT 2021
 - Built deep face counting and deep face tracking pipelines
- **Amazon AWS AI** Seattle, WA
Applied Scientist Intern with Dr. Wei Xia 05/20/2019 - 08/09/2019
 - Designed high resolution face synthesis with disentangled control through facial identity and attributes
 - Presented work to all Amazon Scientist across the globe at Amazon All-Hands meeting
- **Georgia Institute of Technology** Atlanta, GA
Research Assistant with Prof. Jim Rehg 01/09/2017 - 05/04/2018
 - Built a method for gaze target prediction in video
 - Designed a spatial-temporal architecture for gaze target prediction in video
 - Paper [8] published in Computer Vision and Pattern Recognition (CVPR) 2020
 - Participated in designing a method for gaze target prediction in the wild
 - Annotated human gaze targets as in-image and out-of-image for 120,000 sample
 - Performed experiments and designed baseline ablations
 - Published [10] in European Conference of Computer Vision (ECCV) 2018

- **ADP LLC**

Atlanta, GA

Software Development Intern - Full Stack

05/31/2017 - 08/04/2017

- Designed the backend for ADP Web Service Monitor Tool using NodeJS, ExpressJS and MongoDB
- Implemented the front-end user interaction, including user registration, login/logout, customizing service monitoring types

- **Georgia Institute of Technology**

Atlanta, GA

Research Assistant with Prof. John T. Stasko

08/22/2016 - 03/17/2017

- TypoTweet Maps – Characterizing Urban Areas through Typographic Social Media Visualization [11].
 - Designed a novel visualization method to display tweets on as a typographical map with respect to their locations using *d3.js*

PUBLICATIONS

- [1] Y. Xing, X. Tianjun, T. He, **Y. Wang**, Y. Xiong, W. Xia, D.P. Wipf, Z. Zhang, S. Soatto, “Hierarchical graph neural networks for visual clustering”, **US Patent 11860977, 2024** [Details]
- [2] Alexandra Chouldechova, Siqi Deng, **Y. Wang**, Wei Xia, Pietro Perona, “Unsupervised and Semisupervised Bias Benchmarking in Face Recognition”, in **ECCV 2022** [PDF]
- [3] Tianyue Cao , **Y. Wang**, Yifan Xing , Tianjun Xiao , Tong He, Zheng Zhang , Hao Zhou , and Joseph Tighe, “PSS: Progressive Sample Selection for Open-World Visual Representation Learning”, in **ECCV 2022** [PDF]
- [4] Y. Xing, T. He, T. Xiao, **Y. Wang**, Y. Xiong, W. Xia, DW. Paul, Z. Zhang, S. Soatto, “Learning Hierarchical Graph Neural Networks for Image Clustering”, in **ICCV 2021** [PDF] Code
- [5] **Y. Wang***, J. Yang*, R. Yi, Y. Zhu, A. Rehman, A. Zadeh, S. Poria, L.P. Morency, “MTAG: Modal-Temporal Attention Graph for Unaligned Human Multimodal Language Sequences”, in **NAACL-HLT 2021** [PDF] [Code] (* indicates equal contribution)
- [6] **Y. Wang**, K. Kitani, X. Weng, “Joint Object Detection and Multi-Object Tracking with Graph Neural Networks”, in **ICRA 2021** [PDF] [Code (460 Stars/ 73 forks)]
- [7] J. Yang*, Y. Zhu, **Y. Wang**, R. Yi, A. Zadeh, L.P. Morency, “What gives the answer away? question answering bias analysis on video qa datasets”, in **ACL 2020 Human Multimodal Language Workshop** [PDF]
- [8] E. Chong, **Y. Wang**, N. Ruiz, J. Rehg, “Detecting Attended Visual Targets in Video”, in **CVPR 2020** [PDF] [Code (177 Stars/ 48 forks)]
- [9] X. Weng, **Y. Wang**, Y. Man, K. Kitani, “Gnn3dmot: Graph neural network for 3d multi-object tracking with 2d-3d multi-feature learning”, in **CVPR 2020** [PDF]
- [10] E. Chong, N. Ruiz, **Y. Wang**, Y. Zhang, A. Rozga, J. Rehg, “Connecting Gaze, Scene, and Attention: Generalized Attention Estimation via Joint Modeling of Gaze and Scene Saliency”, in **ECCV 2018** [PDF] [Code (177 Stars/ 48 forks)]
- [11] A. Godwin, **Y. Wang**, J. Stasko, “TypoTweet Maps: Characterizing Urban Areas through Typographic Social Media Visualization”, short paper in **European Conference on Visualization 2017** [PDF]

SERVICE

Conference Reviewer for **NeurIPS’21, IJCAI’21, NAACL-HLT’21, ICMI’23**

Journal Reviewer for **IEEE Access**

PROGRAMMING SKILLS

• **Languages:** Python, MatLab, Java, C++ **Technologies:** PyTorch, NumPy, Tensorflow, OpenCV, AWS

SELECTED PROJECTS

- **Face Counting:** Designed a method to count the number of unique faces within a set of images
- **Attended One-Stage Visual Grounding:** Built a project on language grounding in images using attention
- **Face Tracking in Video:** Implemented a project to track human faces within videos while working on Social-IQ
- **Facial-Vocal GAN:** Designed a multi-modal Generative Adversarial Network for talking face generation
- **Multi-Source Domain Adaptation:** Applied Model Agnostic Meta-Learning for multi-source domain adaptation