Jiayi Chen

177 Yellowstone DR APT 308, Charlottesville, VA 22903, United States
jc4td@virginia.edu | (+1) 434-234-7342 | Google Scholar | Linkedin | Homepage

Education

University of Virginia — **United States**
Ph.D. in *Computer Science* (GPA 4.0/4.0) — 08/2018 – present

Xi'an Jiaotong University — **China**
M.S. in *Control Science and Engineering* (Specialization: *Artificial Intelligence and Robotics*) — 09/2015 – 06/2018

Xi'an Jiaotong University — **China**
B.E. in *Automation Engineering* — 09/2011 – 06/2015
Special Class for the Gifted Young of China — 09/2009 – 06/2011

Technical Skills

Programming Languages: **Python** (7+ years), C++, SQL, Java, Matlab
Tools & Platforms: **PyTorch** (5+ years), **Tensorflow**, **Jax** (2+ years), Flax, Haiku, Gym, Robotic Operating System
Deep Neural Networks: Large Language Models (e.g. BERT/ViT/T5/GPT/CLIP/LayoutLM), Graph Neural Networks, Seq2Seq, HyperNetworks, Deep Generative Models (e.g., VAE/GAN/Diffusion), LSTM, …

Research Interests

Multimodal Machine Learning (Vision, Language, Audio, and Spatiotemporal data), Multimodal **Large Language Models**, **Graph** Mining, **Federated** Learning, **Lifelong** Learning, Model Compression (**Pruning/Distillation**)

Recent Publications (Since 2020)

[1] **Jiayi Chen**, Kishlay Jha, Aidong Zhang. "HyperGKL: Diversity-aware Lifelong Learning with Causal Knowledge Hypergraph Learning". *Under Review*, submitted September 2023.

[2] **Jiayi Chen**, Mia Shu, Aidong Zhang. "On Disentanglement of Asymmetrical Knowledge Transfer for Modality-task Agnostic Federated Learning". *Under Review*, submitted August 2023.

[3] **Jiayi Chen**, Hanjun Dai, Bo Dai, Aidong Zhang, Wei Wei. "On Task-personalized Multimodal Few-shot Learning for Visually-rich Document Entity Retrieval". *Findings of Association for Computational Linguistics:* **EMNLP** 2023. (long paper) *Work done during an internship at Google

[4] Lijun Yu, Jin Miao, Xiaoyu Sun, **Jiayi Chen,** AG Hauptmann, Hanjun Dai, Wei Wei. "DocumentNet: Bridging the Data Gap in Document Pre-Training". The *2023 Conference on* **EMNLP**: *Industry Track*. *Work done during internship at Google

[5] **Jiayi Chen**, Aidong Zhang. "On Hierarchical Disentanglement of Interactive Behaviors for Multimodal Spatiotemporal Data with Incompleteness". The *29th ACM SIGKDD Conference on Knowledge Discovery and Data Mining* (**KDD**), 2023. (research track)

[6] **Jiayi Chen**, Aidong Zhang. "FedMSplit: Correlation-adaptive Federated Multitask Learning across Multimodal Split Networks". The *28th ACM SIGKDD Conference on Knowledge Discovery and Data Mining* (**KDD**), 2022. (research track)

[7] Jiachen Yu, Li Jin, **Jiayi Chen,** Youzi Xiao, Zhiqiang Tian, and Xuguang Lan. "Deep Semantic Space guided Multi-scale Neural Style Transfer." *Multimedia Tools and Applications*, 2022, 81(3): 3915-3938. (**Journal**)

[8] **Jiayi Chen**, Aidong Zhang. "Topological Transduction for Hybrid Few-shot Learning". In *Proceedings of the ACM Web Conference* (**WWW**), 2022. (research track)

[9] **Jiayi Chen**, Aidong Zhang. "HetMAML: Task-heterogeneous Model-agnostic Meta-learning for Few-shot Learning across Modalities". *ACM International Conference on Information and Knowledge Management* (**CIKM**), 2021. (full paper)

[10] **Jiayi Chen**, Aidong Zhang. "HGMF: Heterogeneous Graph-based Fusion for Multimodal Data with Incompleteness". The *26th ACM SIGKDD Conference on Knowledge Discovery and Data Mining* (**KDD**), 2020. (research track)

Work Experience

Google DeepMind **Mountain View, CA**
SWE Intern (PhD) June 2023 – Sep 2023

- ***Mentors:*** Zhe Zhao (Google DeepMind), Shuo Yang (YouTube WatchNext), Hussein Hazimeh (Google Research)
- ***Project***: worked on a research project "Topology-agnostic Knowledge Inheritance for Improved Knowledge Distillation." The task is to explore both *direct* and *indirect* knowledge from pre-trained Large Language Models (LLMs) for *recommendation* tasks and vision tasks. The objective is to harness Parameter-Level knowledge of LLMs (e.g., *pre-trained weights* and *their topological structure*) to *accelerate/enhance* knowledge distillation.
- ***Accomplishment:*** a research paper in progress

Google LLC **Mountain View, CA**
SWE Intern (PhD) June 2022 – Dec 2022

- ***Mentors:*** Hanjun Dai (Google DeepMind), Bo Dai (Google DeepMind), Wei Wei (Cloud DocAI & Core ML App)
- ***Project*****:** Efficient Few-shot Multimodal Information Extraction in Visually-rich Documents. Visually-rich documents consist of three modalities (*language*, *image*, and *layout structure* of contents). The task is to harness meta-knowledge to *accelerate* the learning process of (1) understanding *new* document types given a pre-trained ***text-image*** **Large Language Model**; (2) localizing *rarely-occurred* key information types from *out-of-distribution* information.
- ***Accomplishment:*** Successfully published **two** research papers

Project Experience

Understanding Multimodal Data for Semi-/Unsupervised, Few-shot, and Generative Tasks **05/2019 – Present**

- Proposed a graph-factorized **Gaussian Process VAE** to **disentangle** the underlying dynamic systems of partially observed **spatiotemporal interactions** between multiple subjects; and then robustly **generate** imputed data or future data based on learned disentangled representations.
- Proposed a heterogeneous hyper-node **graph** for the **multi-task** multimodal alignment task.
- Proposed a dynamic graph-based **federated learning** approach for **privacy-preserving personalized task adaptation** with heterogeneous domain shifts and modality types.

Toward Emotion Controllable Chatbot via Conditional Generative Models **11/2018 – 03/2019**

- ***Keywords***: NLP, Natural Language Generation, Reinforcement Learning, GANs, VAEs, Sequential Data Generation
- Generate texts/language data whose semantics or attributes (e.g., emotional tones) can be controlled; **adversarially** trained the sequential data generator with Conditional GAN and VAE, while using **reinforcement learning (RL)** to promote the authenticity of the generated language.

Knowledge-guided Neural Style Transfer of 3D Scene Models **03/2017 – 02/2018**

- ***Keywords***: Computer Vision, Neural Style Transfer
- Studied human knowledge-guided **neural style transfer**, focusing on improving the illusion of space in generated images by simulating how artists harness skills to understand and reproduce a 3D scene (e.g., geometric structures, lighting and shallow); also studied 3D non-photorealistic rendering based on the neural style transfer paradigm.
- Proposed an illumination-guided deep alignment method using CNN, Lighting Path Expression, and PatchMatch.
- Created a 2D-3D dataset, including 3D models rendered by multiple types of lighting (by Maya), 2D photos annotated by lighting and segmentation (by Photoshop and Matlab), and a hand-drawn stylistic material for testing (CorelPainter).

Vision-guided Drone Autonomous Navigation & Search-and-rescue System **01/2016 – 12/2016**

- ***Keywords***: Robot Manipulation, Robot Perception, 3D Computer Vision, C++, Robotic Operating System
- Developed a real-time **Vision**-based module (including **Object detection** and **3D localization**) on Drone's autonomous navigation system. Real-time 3D poses of objects are estimated from 2D images using the real-time camera gimbal data based on Homography and 3D Geometry.
- Developed an autonomous **tracking-and-landing** module for Drone landing on a moving vehicle with **safety guarantee**, and designed a platform to manage multimodal sensors and commands (*tools*: **Python**, Linux, ROS).
- Won the **4th** place from 130+ international teams in "2016 DJI Developer Challenge", NY, USA (as a team of three).

TurtleBot Autonomous Security Guard (*course project*) **Fall 2018**

- Built an autonomous framework on TurtleBot to act as a human security guard—wandering, finding AprilTag targets, approaching each target, aiming and then shooting the target with a motorized toy gun installed on TurtleBot.
- Developed the target searching/ranking, goal-position and gun-pitch calculation, and go-to-goal functions (**ROS, python**).

Deep Reinforcement Learning for Human-like Car-Following System (*course project*) **Fall 2020**

- Built an **actor-critic reinforcement learning** framework to learn an optimal car-following behavior from empirical data; implemented **deep deterministic policy gradient** algorithm to learn the continuous-control policy network.

Efficient Human Action Recognition based on Video Compression Domain **Dec. 2014 - May 2015**

- Extracted motion vectors (MV) from MPEG-4 video bitstreams; proposed an MV-based method to **fast** detect Spatial-Temporal Interest Points from video bitstream instead of from the decoded video; formed motion features using BoW and Fisher Vector and trained classifiers like SVM (*tools:* **Matlab, C++, OpenCV, Linux**).

Selected Awards

- ACM KDD'23 Student Travel Award
- **4th** place in world **Top10** in "2016 **International Contest** of DJI Drone Developer Challenge" (*Top 3%, Rank: 4/130*)
- **2nd** prize in "2016 China **National Contest** on Smart-City Technology and Creative Design (3rd Winner Cup)"
- **1st** prize in "2015 China **National Contest** on Traffic-Scene Image Understanding"
- Others: 2018 Outstanding Master's Thesis Award; 2016 Industrial Scholarship (*Top 3%, Rank: 4/104*)

Relevant Courses

Machine Learning and Computer Vision (A+), Natural Language Processing (A+), Reinforcement Learning (A+), Database Systems (A), Graph Mining (A+), Statistical Learning and Graphical Models (A), Computer Graphics (A+), Software Artifacts (A), Autonomous Mobile Robot (A+), Data Structures and Algorithms (A+), Computer Network (A), etc.

Additional

Teaching Assistant: *Program and Data Representation* (Fall 2019, Spring 2020); *Digital Image Processing* (Fall 2015)

Peer Review Services: Reviewer/external reviewer for NeurIPS, ICLR, AAAI, IJCAI, CIKM, WSDM, SDM, BIBM, etc.

Startup Experience: Joined a university **Startup** team (MachInsight) from 2015 to 2016, focusing on dealing with **(1) Computer Vision**-related projects for industry; **(2) Drone** (a type of Autonomous Aerial Robot) related software product driven by Computer Vision; and **(3) Augmented Reality** related product. E.g., at 2015 XJTU University Anniversary, we developed an augmented reality iOS app using Java, Swift, and Unity3D; in 2016, we developed a communicative module for Drone for connecting with Ground Vehicle to achieve enhanced visual tracking and safe landing.