

# Jiayi Chen

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

## Education

<b>University of Virginia</b> Ph.D. Candidate in <i>Computer Science</i> (GPA 4.0/4.0)	<b>United States</b> 08/2018 – 12/2023(Expected)
<b>Xi'an Jiaotong University</b> M.S. in <i>Pattern Recognition &amp; Intelligent System</i> and <i>Control Science &amp; Engineering</i>	<b>China</b> 09/2015 – 06/2018
<b>Xi'an Jiaotong University</b> B. E. in <i>Automation Engineering</i> Special Class for the Gifted Young of China	<b>China</b> 09/2011 – 06/2015 09/2009 – 06/2011

## Research Interests

**2018 – Present:** **Graph** Mining, **Multimodal** Machine Learning (Vision, Language, Audio, Graph, and Spatio-temporal data), **Few-shot** Learning, **Meta** Learning, **Federated** Learning, **Disentangled** Representation Learning, **AutoML**

**2015 – 2018:** Computer Vision, Robotics, Image Processing

## Technical Skills

**Programming Languages:** **Python** (6+ years), C++, Matlab, SQL, HTML, Java

**Tools & Platforms:** **PyTorch** (5+ years), **Tensorflow**, **Jax**, Flax, Haiku, OpenAI Gym, Robotic Operating System

**Deep Neural Networks:** Transformers, Viz Transformers, Graphormers, Large Language Models, Graph Neural Networks, Deep Generative Models (VAEs, GANs, etc.), Encoder-Decoder, Reinforcement Learning models

## Selected Publications (Since 2020)

- [1] **Jiayi Chen**, Mia Shu, Guangzhi Xiong, Aidong Zhang. “HiDenMST: toward hierarchical disentanglement of interactive behaviors for multimodal spatio-temporal data with incompleteness”. Under review, 2023.
- [2] **Jiayi Chen**, Hanjun Dai, Bo Dai, Aidong Zhang, Wei Wei. “On multimodal few-shot learning for visually-rich document entity retrieval”. Under review, 2023.
- [3] **Jiayi Chen**, Mia Shu, Aidong Zhang. “Causality-aware missing-value imputation and future prediction for multivariate time-series”. Under review, 2022.
- [4] **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)
- [5] **Jiayi Chen**, Aidong Zhang. “Adaptive topological transduction for hybrid few-shot learning”. *Proceedings of the International World Wide Web Conference (WWW)*, 2022. (research track)
- [6] **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)
- [7] **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 Experiences

<b>Google LLC</b> Software Engineer Intern (PhD), Cloud <i>Document AI</i> & Google <i>Brain</i> Teams	<b>Mountain View, CA</b> 06/2022 – 09/2022
<ul style="list-style-type: none"><li>• Worked on a <b>research project</b> “<b>Few-shot</b> multimodal entity retrieval from visually rich documents that consists of <i>text</i>, <i>layout structure</i> and <i>image</i> modalities”. The task aims to <i>localize</i> and <i>recognize</i> novel-category entities from new document types given only a few labelled documents. I used <b>meta-learning</b> on the basic of a Large Multimodal <b>Transformer</b> Model encoder to obtain domain-invariant knowledge for the downstream task, which can fast adapt to novel task distributions.</li><li>• Implemented a meta-learning <b>library</b> that covers i) an efficient and balanced task loader; ii) several popular meta-training approaches and a <b>proposed</b> method that outperforms baselines; iii) a bi-level and nested parallel meta-training framework that runs on <b>TPU</b> topology; iv) visualization and evaluation metrics (<i>tools</i>: <b>Jax</b>, <b>Tensorflow</b>, <b>Flax</b>).</li><li>• Tested and documented the designed framework, and ran extensive Xmanager experiments.</li><li>• <i>Accomplishment</i>: a written <b>paper</b> submitted to ACL’23 conference (currently under review)</li></ul>	

## Research & Project Experiences

### On Multi-modality Understanding for Semi-/Unsupervised, Few-shot, or Generative Tasks 05/2019 – Present

- Proposed a graph-factorized **Gaussian Process VAE** to **disentangle** the underlying dynamic systems of partially observed **spatio-temporal dynamics** between multiple subjects; and then robustly **generate** imputed data or future data based on learned disentangled representations.
- Proposed a heterogeneous hypernode **graph** for the **multi-task** multimodal alignment task.
- Proposed a context-aware **meta-learning** approach for fast knowledge distillation from multimodal **few-shot tasks** with heterogeneous domain shifts and modality types.
- Proposed a dynamic graph-based **federated learning** approach for **privacy-preserving personalized task adaptation** with heterogeneous domain shifts and modality types.

### Toward Emotional Chatbot: Natural Language Generation via Conditional Generative Models 11/2018 – 03/2019

- *Keywords:* NLP, Natural Language Generation, Reinforcement Learning, GANs, VAEs, Sequential Data Generation
- Generated sentences 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 generated data.

### 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).

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

- *Keywords:* Deep Reinforcement Learning
- 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.

### 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 on drone system for **object detection** and **3D localization**, computing the 3D poses of objects from 2D images using real-time camera gimbal, Homography and 3D Geometry (*tools:* C++).
- Developed an autonomous **tracking-and-landing** module for landing on a moving vehicle, and developed a platform to manage the interaction between multiple threads of sensors and controllers (*tools:* **Python, Linux, ROS**).
- Achieved safe autonomous navigation during the mission and  $\leq 10\text{cm}$  landing accuracy.
- Won the **4th** place from 130+ international teams in “2016 DJI Developer Challenge”, NY, USA (*as a team of three*).

### Efficient Human Action Recognition based on Video Compression Domain 12/2014 – 05/2015

- *Keywords:* Compressed Video, Human Action Recognition, Video, Spatiotemporal Dynamics, C++
- 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 action features using BoW and GMM; trained traditional classifiers like SVM (*tools:* **Matlab, C++, OpenCV, ffmpeg, Linux**).

## 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.

## Selected Awards

**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** Graduate Contest on Smart-City Technology and Creative Design”

**1st** prize in “2015 China **National** Contest on Traffic-Scene Image Processing”

*Others:* 2022 KDD Student Award; 2018 Outstanding Master’s Thesis Award; 2016 Industrial Scholarship (*Top 3%*)