

# Yapeng TIAN

---

CONTACT INFORMATION 2403 Wegmans Hall 5857669378  
250 Hutchison Road ✉ yapengtian@rochester.edu  
University of Rochester http://yapengtian.org/  
Rochester, NY 14627

RESEARCH INTERESTS My research interests center around solving core **computer vision** and **audition** problems and applying the developed learning approaches to broad AI applications, such as *multisensory perception, computational photography, AR/VR, and HCI*.

RESEARCH AREA Computer Vision Computer Audition Multimodal Learning AI

EDUCATION **University of Rochester**, Rochester, USA Sep. 2017 – Exp. 2022  
• *PhD student* in the Department of Computer Science  
• Advisor: *Prof. Chenliang Xu*  
**Tsinghua University**, Beijing, China Sep. 2014 – July 2017  
• *M.E.* in the Department of Electronic Engineering  
• GPA: 90.55/100 (Rank: 3/52)  
**Xidian University**, Xi'an, China Aug. 2009 – July 2013  
• *B.E.* in Intelligence Science and Technology (School of Electronic Engineering)

WORK EXPERIENCE **Facebook** Sep. 2021 – Jan. 2022  
• *Research Intern* in the Facebook Reality Lab  
• Mentor: *Dr. Alexander Richard*  
**Adobe Research** May 2021 – Aug. 2021  
• *Research Intern* in the Creative Intelligence Lab  
• Mentors: *Dr. Dingzeyu Li* and *Prof. Alexei A. Efros*  
**Adobe Research** May 2019 – Nov. 2019  
• *Research Intern* in the Creative Intelligence Lab  
• Mentor: *Dr. Dingzeyu Li*

RESEARCH EXPERIENCE **CS, University of Rochester** Aug. 2017 – Present  
• *Research Assistant* with *Prof. Chenliang Xu*  
**EE, Tsinghua University** Mar. 2015 – Aug. 2017  
• *Research Assistant* with *Prof. Wenming Yang*  
**SIAT, Chinese Academy of Sciences** Nov. 2016 – May 2017  
• *Visiting Student* with *Prof. Yu Qiao*

## PUBLICATIONS

*CVPR, ICCV, and ECCV are premier computer vision conferences. According to Google Scholar Metrics, as of 3/2022, CVPR has h5-index 356, ECCV 197, and ICCV 184. CVPR is also ranked 1st of all journals and conferences in Engineering and Computer Science and 4th when considering everything else. Citations: 4317, h-index: 13, i10-index: 16 by Google Scholar, 3/2022.*

### Conference Papers (7 CVPR, 2 ECCV, 2 ICCV, 2 AAAI, 1 BMVC)

- Guangyao Li\*, Yake Wei\*, **Yapeng Tian\***, Chenliang Xu, Ji-Rong Wen, and Di Hu. Learning to Answer Questions in Dynamic Audio-Visual Scenarios. *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 2022. (**\*Equal contribution, Oral**)
- Guangyuan Li, Jun Lv, **Yapeng Tian**, Qi Dou, Chengyan Wang, Chenliang Xu, and Jing Qin. Transformer-empowered Multi-scale Contextual Matching and Aggregation for Multi-contrast MRI Super-resolution. *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 2022.
- Bin Xia, **Yapeng Tian**, Yucheng Hang, Wenming Yang, Qingmin Liao, and Jie Zhou. Coarse-to-Fine Embedded PatchMatch and Multi-Scale Dynamic Aggregation for Reference-based Super-Resolution. *The AAAI Conference on Artificial Intelligence (AAAI)*, 2022.
- Bin Xia\*, Yucheng Hang\*, **Yapeng Tian**, Wenming Yang, Qingmin Liao, and Jie Zhou. Efficient Non-Local Contrastive Attention for Image Super-Resolution. *The AAAI Conference on Artificial Intelligence (AAAI)*, 2022. (**\*Equal contribution.**)
- Sizhe Li\*, **Yapeng Tian\***, and Chenliang Xu. Space-Time Memory Network for Sounding Object Localization in Videos. *The British Machine Vision Conference (BMVC)*, 2021. (**\*Equal contribution.**)
- Tiantian Wang, Sifei Liu, **Yapeng Tian**, Kai Li, and Ming-Hsuan Yang. Video Matting via Consistency-Regularized Graph Neural Networks. *IEEE/CVF International Conference on Computer Vision (ICCV)*, 2021.
- **Yapeng Tian**, and Chenliang Xu. Can audio-visual integration strengthen robustness under multimodal attacks? *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 2021.
- **Yapeng Tian**, Di Hu, and Chenliang Xu. Cyclic Co-Learning of Sounding Object Visual Grounding and Sound Separation. *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 2021.
- **Yapeng Tian**, Dingzeyu Li, and Chenliang Xu. Unified Multisensory Perception: Weakly-Supervised Audio-Visual Video Parsing. *European Conference on Computer Vision (ECCV)*, 2020. (**Spotlight, top 5%**)
- **Yapeng Tian**, Yulun Zhang, Yun Fu, and Chenliang Xu. TDAN: Temporally-Deformable Alignment Network for Video Super-Resolution. *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 2020.
- Xiaoyu Xiang\*, **Yapeng Tian\***, Yulun Zhang, Yun Fu, Jan Allebach, and Chenliang Xu. Zooming Slow-Mo: Fast and Accurate One-Stage Space-Time Video Super-Resolution. *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 2020. (**\*Equal contribution.**)
- Wei Wang\*, Ruiming Guo\*, **Yapeng Tian**, and Wenming Yang. CFSNet: Toward a

Controllable Feature Space for Image Restoration. *IEEE/CVF International Conference on Computer Vision (ICCV)*, 2019. (*\*Equal contribution.*)

- **Yapeng Tian**, Jing Shi, Bochen Li, Zhiyao Duan, and Chenliang Xu. Audio-Visual Event Localization in Unconstrained Videos. *European Conference on Computer Vision (ECCV)*, 2018.
- Yulun Zhang, **Yapeng Tian**, Yu Kong, Bineng Zhong, Yun Fu. Residual Dense Network for Image Super-Resolution. *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 2018. (**Spotlight**, top 5%)
- Xuesen Shang, Wenming Yang, Shuifa Sun, **Yapeng Tian**, Hai Chen, Kaiquan Chen. Adaptive Anchor-Point Selection for Single Image Super-Resolution. *IEEE International Conference on Visual Communications and Image Processing (VCIP)*, 2017.
- **Yapeng Tian**, Fei Zhou, Wenming Yang, Xuesen Shang and Qingmin Liao. Anchored Neighborhood Regression based Single Image Super-Resolution from Self-Examples. *IEEE International Conference on Image Processing (ICIP)*, 2016.
- Wenming Yang, **Yapeng Tian**, Fei Zhou, Tingrong Yuan, Xuesen Shang and Qingmin Liao. Single-Image Super-Resolution Using Clustering-Based Global Regression and Propagation Filtering. *Asian Conference on Pattern Recognition (ACPR)*, 2015. (**Oral**, top 8%)

#### Journal Papers (1 TPAMI, 1 TIP, 2 TMM)

- Yulun Zhang, **Yapeng Tian**, Yu Kong, Bineng Zhong, Yun Fu. Residual Dense Network for Image Restoration. *IEEE Transactions on Pattern Analysis and Machine Intelligence (TPAMI)*, 2020.
- Wenming Yang, Xuechen Zhang, **Yapeng Tian**, Wei Wang, Jing-Hao Xue, Qingmin Liao. LCSCNet: Linear Compressing Based Skip-Connecting Network for Image Super-Resolution. *IEEE Transactions on Image Processing (TIP)*, 2019.
- Wenming Yang, Xuechen Zhang, **Yapeng Tian**, Wei Wang, Jing-Hao Xue, Qingmin Liao. Deep Learning for Single Image Super-Resolution: A Brief Review. *IEEE Transactions on Multimedia (TMM)*, 2019.
- Wenming Yang, **Yapeng Tian**, Fei Zhou, Qingmin Liao, Hai Chen and Chenglin Zheng. Consistent Coding Scheme for Single-Image Super-Resolution Via Independent Dictionaries. *IEEE Transactions on Multimedia (TMM)*, 2016.

#### Workshop Papers (6 CVPR Workshop)

- **Yapeng Tian**, Chenliang Xu, Dingzeyu Li. Deep Audio Prior: Learning Sound Source Separation from a Single Audio Mixture. *IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPR Workshop)*, 2020.
- **Yapeng Tian\***, Di Hu\*, Chenliang Xu. Co-Learn Sounding Object Visual Grounding and Visually Indicated Sound Separation in A Cycle. *IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPR Workshop)*, 2020. (*\*Equal contribution.*)
- **Yapeng Tian**, Dingzeyu Li, and Chenliang Xu. Weakly-Supervised Audio-Visual Video Parsing Toward Unified Multisensory Perception. *IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPR Workshop)*, 2020.
- **Yapeng Tian**, Chenxiao Guan, Goodman Justin, Marc Moore, and Chenliang Xu.

Audio-Visual Interpretable and Controllable Video Captioning. *IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPR Workshop)*, 2019.

- **Yapeng Tian**, Jing Shi, Bochen Li, Zhiyao Duan, and Chenliang Xu. Audio-Visual Event Localization in the Wild. *IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPR Workshop)*, 2019. (**Oral**, 1/10)
- Timofte *et al.* NTIRE 2017 Challenge on Single Image Super-Resolution: Methods and Results. *IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPR Workshop)*, 2017.

## WORK IN PROGRESS

- Zheng Zhang\*, **Yapeng Tian\***, Zheng Ning, Chenliang Xu, and Toby Jia-Jun Li. PEANUT: An Intelligent Human-AI Collaborative Tool for Annotating Audio-Visual Data, *Work in Progress*. (**\*Equal contribution.**)
- **Yapeng Tian**, Alexei A. Efros, Chenliang Xu, and Dingzeyu Li. HelpDescribe: Accelerating Audio Description Creation with Human-in-the-loop Recommendation, *Work in Progress*.
- Yiyang Su\*, **Yapeng Tian\***, and Chenliang Xu. Separating Invisible Sounds toward Universal Audio-Visual Scene-Aware Sound Separation, *Work in Progress*. (**\*Equal contribution.**)
- Rohan Sharma\*, **Yapeng Tian\***, and Chenliang Xu. Cooperative Audio-Visual Video Parsing and Captioning, *Work in Progress*. (**\*Equal contribution.**)

## TEACHING EXPERIENCE

### Teaching Assistant

- *Machine Vision*, University of Rochester Spring 2019
- *Advanced Topics in Computer Vision*, University of Rochester Fall 2018
- *Machine Vision*, University of Rochester Spring 2018
- *Advanced Image Processing and Its Applications*, Tsinghua University Fall 2016
- *Digital Image Processing*, Tsinghua University Spring 2016

### Guest Lecturer

- *Advanced Topics in Computer Vision*, University of Rochester Spring 2021
- *Machine Vision*, University of Rochester Fall 2020

## HONORS AND AWARDS

- Top 10% of High-Scoring Reviewers for NeurIPS 2020
- Invited attendee of Amazon Graduate Student Symposium, Seattle, USA 2019
- Outstanding Graduate of Tsinghua University (Top 1%) 2017
- Outstanding Master Thesis Award, Tsinghua University 2017
- National Scholarship, Tsinghua University (Top 2%) 2016
- Second-class Scholarship, Tsinghua University 2015

## PROFESSIONAL ACTIVITIES

### Organizing Committee

- *Audio-Visual Scene Understanding Tutorial* at CVPR June 2021
- *Audio-Visual Scene Understanding Tutorial* at WACV Jan. 2021

### Talks and Seminars

- *Audio-Visual Scene Understanding*  
*Towards Unified, Explainable, and Robust Multisensory Perception*  
University of Texas at Dallas Mar. 2022  
George Mason University Feb. 2022  
Dartmouth College Feb. 2022  
KTH Dive-Deep Seminar Dec. 2021  
RIT PhD Colloquium Series Oct. 2021
- *Audio-Visual Video Understanding*, IIAI Seminar Sep. 2021
- *The Future of Audio-Visual Research Panel Discussion*, VALSE Webinar Nov. 2020

### Program Committee/Conference Reviewer

- CVPR: IEEE/CVF Conference on Computer Vision and Pattern Recognition 2019-2022
- ICCV: IEEE/CVF International Conference on Computer Vision 2019-2021
- ECCV: European Conference on Computer Vision 2020-2022
- NeurIPS: Conference on Neural Information Processing Systems 2020-2021
- ICLR: International Conference on Learning Representations 2021-2022
- AAAI: AAAI Conference on Artificial Intelligence 2020-2022
- ICML: International Conference on Machine Learning 2021-2022
- WACV: Winter Conference on Applications of Computer Vision 2020-2021
- ACCV: Asian Conference on Computer Vision 2021

### Journal Reviewer

- TPAMI: IEEE Transactions on Pattern Analysis and Machine Intelligence 2021
- TMLR: The Transactions on Machine Learning Research 2021-2022
- TIP: IEEE Transactions on Image Processing 2021-2022
- TNNLS: IEEE Transactions on Neural Networks and Learning Systems 2021
- TMM: IEEE Transactions on Multimedia 2019-2022
- TCSVT: IEEE Transactions on Circuits and Systems for Video Technology 2019-2022
- Scientific Reports – Nature 2021
- IEEE Access 2019-2021
- SPIC: Signal Processing: Image Communication 2018-2021
- CVIU: Computer Vision and Image Understanding 2020
- CGF: Computer Graphics Forum 2020

### UNIVERSITY SERVICES

- CS PhD Admission's Committee, University of Rochester 2018-2022
- ASE Conference Travel Funding Grant Reviewer, University of Rochester 2018

### STUDENT ADVISING

#### MS Students

- Rohan Sharma (Data Science, UofR → PhD student at SUNY Buffalo) 2020-2021  
Project: *audio-visual scene-aware captioning*
- Purvanshi Mehta (Data Science, UofR → Data Scientist at Microsoft) Spring 2020

Project: *multimodal continual learning*

- Shurui Zhang (Optics, UofR) 2018  
Project: *video super-resolution*

#### Undergraduates

- Sizhe Li (Computer Science, UofR → Research Intern at MIT CSAIL) 2019–2021  
Project: *sounding object visual localization* → *BMVC 2021*
- Yiyang Su (Computer Science, UofR → PhD student at MSU) 2020–2021  
Project: *separating invisible sounds*
- Chenxiao Guan (Xerox Fellow at UofR → Master student at CMU) Summer 2018  
Project: *audio-visual video captioning* → *CVPR Workshop 2019*
- Justin Goodman (UMD, REU at UofR → Master student at UMD) Summer 2018  
Project: *audio-visual data collection* → *CVPR Workshop 2019*
- Marc Moore (Mississippi State University, REU at UofR) Summer 2018  
Project: *audio-visual data collection* → *CVPR Workshop 2019*

#### SKILLS

- Languages: English, Mandarin (native).
- Programming: Python, Pytorch, Keras, MATLAB,  $\LaTeX$ .