Qi Chen

Homepage · Google Scholar · Citations>1,100 · H-index=14

Linkedin Profile · chenqigg123@gmail.com · qi.chen04@adelaide.edu.au

(+61) 493411360 (in Australia) · (+86) 188-141-22530 (in China)

About Me

Dr. Qi Chen is currently a postdoctoral research fellow at the Australian Institute for Machine Learning (AIML), the University of Adelaide, working with *Prof. Anton van den Hengel* and *A/Prof. Qi Wu*. He obtained his PhD in 2024 from the University of Adelaide (92th in 2024 U.S. News), under the supervision of *A/Prof. Qi Wu* and *Asst. Prof. Yuankai Qi*. He received his Master's and Bachelor's degrees from South China University of Technology (SCUT) in China (187th in 2024 U.S. News) under the supervision of *Prof. Jian Chen* and *Prof. Mingkui Tan*.

Dr. Chen focuses mainly on Controllable Generative AI for Multi-modality, (Multimodal) Large Language Models (LLMs), and Multimodal AI for Real-world Applications/Domains (e.g., Medicine, Architecture, and the Internet). He has over 20 peer-reviewed publications, most in flagship journals/conference proceedings, including IEEE-TPAMI/TIP/TMM, CVPR, NeurIPS, ICCV, etc. His research has attracted over 1,100 citations with an H-index of 14 (Google Scholar). He also serves as a reviewer for top-tier journals/conference proceedings, including Nature Communications, IEEE-TPAMI, IJCV, CVPR, ICML, NeurIPS, ICLR, ICCV, ECCV, etc.

Education

The University of Adelaide, Australia 2021 – 2024

PhD candidate at the School of Computer Science, Supervisor: *A/Prof. Qi Wu* & *Asst. Prof. Yuankai Qi*

South China University of Technology (SCUT), China 2017 – 2020

Master student at the School of Software Engineering, Supervisor: *Prof. Jian Chen* & *Prof. Mingkui Tan*

South China University of Technology (SCUT), China 2013 – 2017

Bachelor student at the School of Software Engineering

Internship and Work Experience

Postdoctoral Researcher 2024 – Present

Australian Institute for Machine Learning (AIML), The University of Adelaide
Project: Close the Gap between Machine and Human
Work with *Prof. Anton van den Hengel* and *A/Prof. Qi Wu*

Research Assistant 2020 – 2022

Guangdong Artificial Intelligence and Digital Economy Laboratory (Pazhou Lab)
Project: Fundamental Theories and Core Algorithms of AI
Work with *Prof. Yuanqing Li* (IEEE Fellow) & *Prof. Mingkui Tan*

Nine Best Publications (* equal contributions)

Controllable Generative AI for Multi-modality

Enhance the controllability of generative AI by broadening modality coverage (text, image, audio, etc), extending dimensions (e.g., from static images to dynamic videos), and refining granularity (e.g., linguistic style-aware image captioning).

1. **Learning Distinct and Representative Modes for Image Captioning**
   Qi Chen, Chaorui Deng, Qi Wu

Advances in Neural Information Processing Systems (NeurIPS), 2022. (Top-tier Conference on Machine Learning)

2. **V2C: Visual Voice Cloning**
   Qi Chen, Mingkui Tan, Yuankai Qi, Jiaqiu Zhou, Yuanqing Li, Qi Wu
   IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2022. (Top-tier Conference on Computer Vision)
3. **Scripted Video Generation with a Bottom-up Generative Adversarial Network**
   Qi Chen, Qi Wu, Jian Chen, Qingyao Wu, Anton van den Hengel, Mingkui Tan
   IEEE Transactions on Image Processing (TIP), 2020. (Impact factor=10.8, Top-tier Journal on Computer Vision & Pattern Recognition)
4. **G-NeRF: Geometry-enhanced Novel View Synthesis from Single-View Images**
   Zixiong Huang*, **Qi Chen***, Libo Sun, Yifan Yang, Naizhou Wang, Mingkui Tan, Qi Wu
   IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2024. (Top-tier Conference on Computer Vision)

(Multimodal) Large Language Models (LLMs)

Integrate and interpret information from diverse data sources, enable more comprehensive and context-aware analyses, allow for a deeper understanding of complex scenarios, assess the quality of a new AI by (multimodal) LLMs, and then enable more accurate and effective reasoning and decision-making.

1. **Weak-eval-Strong: Evaluating and Eliciting Lateral Thinking of LLMs with Situation Puzzles**
   Qi Chen, Bowen Zhang, Gang Wang, Qi Wu
   Advances in Neural Information Processing Systems (NeurIPS), 2024. (Top-tier Conference on Machine Learning)
2. **Prompt Switch: Efficient CLIP Adaptation for Text-Video Retrieval**
   Chaorui Deng*, **Qi Chen***, Pengda Qin, Da Chen, Qi Wu
   International Conference on Computer Vision (ICCV), 2023. (Top-tier Conference on Computer Vision)

Multimodal AI for Real-world Applications/Domains

Span multimodal AI across various domains, including architectural design, where it aids in visualizing and optimizing space usage; medical diagnostics, where it enhances patient care by comprehensive data analysis; website assistance, where it improves user interaction by intuitively responding to both text and voice queries.

1. **Intelligent Home 3D: Automatic 3D-House Design from Linguistic Descriptions Only**
   Qi Chen, Qi Wu, Rui Tang, Yuhan Wang, Shuai Wang, Mingkui Tan
   IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2020. (Top-tier Conference on Computer Vision) (Google citation: >35)
2. **WebVLN: Vision-and-Language Navigation on Websites**
   Qi Chen, Dileepa Pitawela, Chongyang Zhao, Gengze Zhou, Hsiang-Ting Chen, Qi Wu
   Association for the Advancement of Artificial Intelligence (AAAI), 2024. (Top-tier Conference on Artificial Intelligence)
3. **PairAug: What Can Augmented Image-Text Pairs Do for Radiology?**
   Yutong Xie*, **Qi Chen***, Sinuo Wang, Minh-Son To, Iris Lee, Ee Win Khoo, Kerolos Hendy, Daniel Koh, Yong Xia, Qi Wu
   IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2024. (Top-tier Conference on Computer Vision)

🕮 Other Publications (* equal contributions; † corresponding author)

Peer-refereed Conference Proceedings

1. **Act Like a Radiologist: Radiology Report Generation across Anatomical Regions**
   Qi Chen, Yutong Xie, Biao Wu, Xiaomin Chen, James Ang, Minh-Son To, Xiaojun Chang, Qi Wu
   Proceedings of the Asian Conference on Computer Vision (ACCV), 2024. (Oral, 17.4% of accepted papers, 5.6% of submitted papers)
2. **R-GAN: Exploring Human-like Way for Reasonable Text-to-Image Synthesis via Generative Adversarial Networks**

Yanyuan Qiao, **Qi Chen**, Chaorui Deng, Ning Ding, Yuankai Qi, Mingkui Tan, Xincheng Ren, Qi Wu
ACM International Conference on Multimedia (ACM MM), 2021. (Top-tier Conference on Multimedia)

3. **Contrastive Neural Architecture Search with Neural Architecture Comparators**
Yaofo Chen, Yong Guo, **Qi Chen**, Minli Li, Yaowei Wang, Wei Zeng, Mingkui Tan
IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2021. (Top-tier Conference on Computer Vision) (Google citation: >70)
4. **Closed-loop Matters: Dual Regression Networks for Single Image Super-Resolution**
Yong Guo, Jian Chen, Jingdong Wang, **Qi Chen**, Jiezhang Cao, Zeshuai Deng, Yanwu Xu, Mingkui Tan
IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2020. (Top-tier Conference on Computer Vision) (Google citation: >400)
5. **Dynamic Extension Nets for Few-shot Semantic Segmentation**
Lizhao Liu*, Junyi Cao*, Minqian Liu*, Yong Guo*, **Qi Chen***, Mingkui Tan
ACM International Conference on Multimedia (ACM MM), 2020. (Top-tier Conference on Multimedia) (Google citation: >50)
6. **NAT: Neural Architecture Transformer for Accurate and Compact Architectures**
Yong Guo, Yin Zheng, Mingkui Tan, **Qi Chen**, Jian Chen, Peilin Zhao, Junzhou Huang
Advances in Neural Information Processing Systems (NeurIPS), 2019. (Top-tier Conference on Machine Learning) (Google citation: >90)
7. **Modular graph attention network for complex visual relational reasoning**
Yihan Zheng, Zhiquan Wen, Mingkui Tan, Runhao Zeng, **Qi Chen**, Yaowei Wang, Qi Wu
Proceedings of the Asian Conference on Computer Vision (ACCV), 2020.

Peer-refereed Journal Publications

1. **Auto-embedding Generative Adversarial Networks for High Resolution Image Synthesis**
Yong Guo*, **Qi Chen***, Jian Chen, Qingyao Wu, Qinfeng Shi, Mingkui Tan
IEEE Transactions on Multimedia (TMM), 2019. (Impact factor=8.4, Top-tier Journal on Multimedia) (Google citation: >80)
2. **Auto-3D-House Design from Structured User Requirements**
Mingkui Tan*, **Qi Chen***, Zixiong Huang, Qi Wu, Yuanqing Li, Jiaqiu Zhou
Machine Intelligence Research (MIR), 2024. (Early Access) (Impact factor=6.4)
3. **Towards Accurate and Compact Architectures via Neural Architecture Transformer**
Yong Guo, Yin Zheng, Mingkui Tan, **Qi Chen**, Zhipeng Li, Jian Chen, Peilin Zhao, Junzhou Huang
IEEE Transactions on Pattern Analysis and Machine Intelligence (TPAMI), 2021. (Impact factor=20.8, Top-tier Journal on Computer Vision & Pattern Recognition) (Google citation: >35)
4. **Towards Lightweight Super-Resolution with Dual Regression Learning**
Yong Guo, Mingkui Tan, Zeshuai Deng, Jingdong Wang, **Qi Chen**, Jiezhang Cao, Yanwu Xu, Jian Chen
IEEE Transactions on Pattern Analysis and Machine Intelligence (TPAMI), 2024. (Impact factor=20.8, Top-tier Journal on Computer Vision & Pattern Recognition)

Preprints

1. **Likelihood-Based Text-to-Image Evaluation with Patch-Level Perceptual and Semantic Credit Assignment**
Qi Chen, Chaorui Deng, Zixiong Huang, Bowen Zhang, Mingkui Tan, Qi Wu
arXiv:2308.08525
2. **Attention-guided Context Feature Pyramid Network for Object Detection**
Junxu Cao*, **Qi Chen***, Jun Guo, Ruichao Shiu
arXiv:2005.11475 (Google citation: >130)
3. **Efficient Response Selection for Fine-Tuning Large Language Models via Alignment Score Estimation**
Xuan Ren*, **Qi Chen***, Lingqiao Liu
4. **Dual Reconstruction Nets for Image Super-resolution with Gradient Sensitive Loss**
Yong Guo, **Qi Chen**, Jian Chen, Junzhou Huang, Yanwu Xu, Jiezhang Cao, Peilin Zhao, Mingkui Tan
arXiv:1809.07099
5. **InfiniMotion: Mamba Boosts Memory in Transformer for Arbitrary Long Motion Generation**
Zeyu Zhang, Akide Liu, **Qi Chen**, Feng Chen, Ian Reid, Richard Hartley, Bohan Zhuang, Hao Tang

arXiv:2407.10061

6. **Source-Free Unsupervised Domain Adaptation with Hypothesis Consolidation of Prediction Rationale**
   Yangyang Shu, Xiaofeng Cao, **Qi Chen**, Bowen Zhang, Ziqin Zhou, Anton van den Hengel, Lingqiao Liu
   arXiv:2402.01157
7. **Xlip: Cross-modal Attention Masked Modelling for Medical Language-image Pre-training**
   Biao Wu, Yutong Xie, Zeyu Zhang, Minh Hieu Phan, **Qi Chen**, Ling Chen, Qi Wu
   arXiv:2407.19546
8. **A Thorough Comparison Study on Adversarial Attacks and Defenses for Common Thorax Disease Classification in Chest X-rays**
   Chendi Rao, Jiezhang Cao, Runhao Zeng, **Qi Chen**, Huazhu Fu, Yanwu Xu, Mingkui Tan
   arXiv:2003.13969

Chinese Journal Publications and Patents

1. **Large Language Model-guided Video-text Retrieval Data Optimization**
   Runhao Zeng, Jialiang Li, Yishen Zhuo, Haihan Duan, **Qi Chen**[†], Xiping Hu
   Journal of Image and Graphics, 2024
2. **A Review on Deep Adversarial Visual Generation**
   Mingkui Tan, Shoukai Xu, Shuhai Zhang, **Qi Chen**
   Journal of Image and Graphics, 2021
3. **Quantitative Investment Trend Prediction System Based on Deep Learning Algorithms**
   Qi Chen, Jian Chen
   2018SR766183, published

♡ Honors and Awards

Research Grants

National Natural Science Foundation of China (about A$117,500) (CI) 2021 – 2024

Awards

Australian Government-funded Research Training Program (RTP) Scholarship
($\sim$ 300 international students globally each year) 2021 – 2024

Golden Award of CICSIC Challenge (International Track)
(Awarded to 50 teams, top 0.1% globally) 2024

First-class Scholarship of South China University of Technology (SCUT)
(Top 1% students, annually) 2017 – 2020

⚙ Professional Activities

Conference Reviews

[Computer Vision]

CVPR 2020-2024, ICCV 2021/2023, ECCV 2020/2022/2024, WACV 2022-2025, BMVC 2022/2023, ACCV 2022/2024

[Machine Learning]

NeurIPS 2022-2024, ICML 2022-2024, ICLR 2024/2025, AISTATS 2025

[Others]

AAAI 2020-2024, MICCAI 2024, SIGGRAPH2024

Journal Reviews

Nature Communications

IEEE Transactions on Pattern Analysis and Machine Intelligence (TPAMI)

International Journal of Computer Vision (IJCV)

Pattern Recognition (PR), etc.

Presentations

Weak-eval-Strong: Evaluating and Eliciting Lateral Thinking of LLMs with Situation Puzzles
Venue: NeurIPS 2024 (coming soon) Dec 2024

Act Like a Radiologist: Radiology Report Generation across Anatomical Regions
Venue: ACCV 2024 Oral Presentation (coming soon) Dec 2024

G-NeRF: Geometry-enhanced Novel View Synthesis from Single-View Images
Venue: AIML Showcase event, the National Wine Centre, Adelaide Oct 2024

Webvln: Vision-and-Language Navigation on Websites
Venue: AAAI 2024 Feb 2024

Multimodal Approaches and Applications in Vision, Language and Audio
Venue: 2023 Global Young Scholars' Forum at CUHK-Shenzhen Dec 2023

Prompt Switch: Efficient CLIP Adaptation for Text-Video Retrieval
Venue: ICCV 2023 Oct 2023

Learning Distinct and Representative Modes for Image Captioning
Venue: AIML Showcase event, Adelaide Oval Aug 2022

Intelligent Home 3D: Automatic 3D-House Design from Linguistic Descriptions Only
Venue: CSIG-Guangdong Province CVPR 2020 Online Academic Seminar May 2020

NAT: Neural Architecture Transformer for Accurate and Compact Architectures
Venue: NeurIPS 2019 Dec 2019

Teaching Experience

Teaching Assistant 2023 Semester 1

Course: Advanced Topics in Computer Science
School of Computer and Mathematical Sciences, The University of Adelaide

Teaching Assistant 2022 Trimester 2

Course: Artificial Intelligence and Machine Learning Research Project Part B
School of Computer and Mathematical Sciences, The University of Adelaide

Teaching Assistant 2022 Trimester 1

Course: Artificial Intelligence and Machine Learning Research Project Part A
School of Computer and Mathematical Sciences, The University of Adelaide

Student Supervision

Jinquan Guan (Ph.D. student) of SCUT
Achievement: One paper has been submitted to International Journal of Surgery
Position: Co-supervisor, Instructor: Prof. Jian Chen Sep 2024 – present

Zixiong Huang (Master student) of SCUT
Achievement: One paper has been accepted by CVPR 2024
Position: Co-supervisor, Instructor: Prof. Mingkui Tan Sep 2022 – present

Xiaomin Chen (Master student) of SCUT
Achievement: One paper has been accepted by NeurIPS 2024 workshop
Position: Co-supervisor, Instructor: Prof. Jian Chen Sep 2022 – present

Jialiang Li (Master student) of Shenzhen MSU-BIT University
Achievement: One paper has been accepted by Journal of Image and Graphics 2024
Position: Co-supervisor, Instructor: A/Prof. Runhao Zeng Sep 2022 – present

Yishen Zhuo (Master student) of Shenzhen MSU-BIT University
Achievement: One paper has been submitted to TCSVT
Position: Co-supervisor, Instructor: A/Prof. Runhao Zeng Sep 2022 – present

Jiaqiu Zhou (Master student) of SCUT
Achievement: One paper has been accepted by Machine Intelligence Research
Position: Co-supervisor, Instructor: Prof. Mingkui Tan Sep 2019 – Jun 2022

REFERENCES

1. Anton van den Hengel, Professor, University of Adelaide; anton.vandenhengel@adelaide.edu.au
2. Qi Wu, Associate Professor, University of Adelaide; qi.wu01@adelaide.edu.au
3. Mingkui Tan, Professor, South China University of Technology; mingkuitan@scut.edu.cn
4. Yuankai Qi, Assistant Professor, Macquarie University; qykshr@gmail.com
5. Jian Chen, Professor, South China University of Technology; ellachen@scut.edu.cn
6. Yutong Xie, Assistant Professor, MBZUAI; yutong.xie678@gmail.com