

XIAOQIAN SHEN

[Personal page](#) / [Github](#)

Email xiaoqian.shen@kaust.edu.sa

Tel: (+966)542708372 / (+86)13642093218

EDUCATION

Master of Computer Science KAUST Expected 2024
King Abdullah University of Science and Technology

Bachelor of Computer Science Jilin University 2018 - 2022
GPA: 90.27/100 (3.73/4.0)

RESEARCH INTEREST

- **Generative Models**
- **Vision-Language Multimodal Learning**
- **Spatiotemporal Representation**

PROJECTS

Master at Prof. Mohamed Elhoseiny's group, KAUST, Saudi Arabia

MiniGPT-4 April 2023
Uncover the recipe and reproduce the multi-modal capacity of GPT-4. [\[project page\]](#)

ChatCaptioner Feb 2023 - Mar 2023
Enriched visual description through interaction between vision and language foundation model.

Motion-aware Efficient Video Generation July 2022 - Nov 2022
Propose motion style attention modulation for diverse motion pattern modeling. [\[project page\]](#)

Affective Visual Dialog July 2022 - Nov 2022
Construct a VQA dataset to analyze how visual and language affect emotion. [\[project page\]](#)

Implicit Text to Image Generation Aug 2022 - Sep 2022
Integrate conditional signals to implicit neural representations (INR)-based image generation. [\[project page\]](#)

Internship at Prof. Mohamed Elhoseiny's group, KAUST, Saudi Arabia

Large-scale Zero-shot Classification Dec 2021 - Mar 2022
Leverage hierarchical contrastive learning for large-scale zero-shot classification. [\[project page\]](#)

Internship at Prof. Yongfeng Huang's group, Tsinghua University, China

Medical Relation Extraction for Chinese Medicine Instructions Sep 2020 - Mar 2021
Explore BERT-like model for text modeling and entity recognition.

Internship at Prof. Zhenghua Xu's group, External D.Phil Supervisor of University of Oxford

Self-supervised Learning for medical image segmentation May 2020 - Sep 2020
Utilize multimodality information of medical images for downstream tumor segmentation.

PUBLICATIONS

Xiaoqian Shen, Xiang Li, and Mohamed Elhoseiny. MoStGAN-V: Video Generation with Temporal Motion Styles. CVPR 2023. [\[paper\]](#)

Deyao Zhu, Chen Jun, **Xiaoqian Shen**, Xiang Li, Mohamed Elhoseiny. MiniGPT-4: Enhancing Vision-language Understanding with Advanced Large Language Models. Arxiv. [\[paper\]](#)

Deyao Zhu, Chen Jun, Kilichbek Haydarov, **Xiaoqian Shen**, Wenxuan Zhang, Mohamed Elhoseiny. ChatGPT Asks, BLIP-2 Answers: Automatic Questioning Towards Enriched Visual Descriptions. Arxiv. [\[paper\]](#)

Eslam Bakr, **Xiaoqian Shen**, Pengzhan Sun, Faizan Khan, Erran Li, Mohamed Elhoseiny. HRS-Bench: Holistic, Reliable and Scalable Benchmark for Text-to-Image Models. Arxiv. [[paper](#)]

Yi, Kai, **Xiaoqian Shen**, Yunhao Gou, and Mohamed Elhoseiny. Exploring hierarchical graph representation for large-scale zero-shot image classification. Computer Vision–ECCV 2022, pp. 116-132. [[paper](#)]

Qi, Tao, Shan Qiu, **Xiaoqian Shen**, Haopu Chen, Shuai Yang, Hao Wen, Ya Zhang, Yuanqing Wu, and Yongfeng Huang. KeMRE: knowledge-enhanced medical relation extraction for Chinese medicine instructions. Journal of Biomedical Informatics 120 (2021): 103834. [[paper](#)]

SKILLS

- **Programming:** Python, C/C++, Java, HTML
- **Language** Chinese, English (TOEFL 104/120, GRE 328/340)