Tong Zhang

tongz27@uci.edu +1 949-232-5050 the-star-sea Homepage

Education Background

University of California Irvine (UCI) **Irvine, USA**
MS, Computer Engineering, GPA: ongoing *June 2023 - June 2024*

Southern University of Science and Technology (SUSTech) **Shenzhen, China**
Bachelor of Computer Science and Engineering, GPA: 3.65/4 *Sep. 2019 - June 2023*

University of California Irvine (UCI) **Irvine, USA**
Semester Exchange, GPA: 3.62/4 *June 2022 - June 2023*

Research Interest

My research interest lies at the intersection of multimodal learning and trustworthy machine learning, focusing on developing intelligent systems that enhance human-related visual understanding and generation. This involves several key areas:

- **Multimodal Data Interpretation:** Bridging the gap between human cognitive processes and machine learning, particularly in interpreting complex data like images, texts, and videos to create coherent and meaningful representations.
- **SVG Generation and Interpretation:** Developing methodologies to convert complex image data into scalable vector graphics (SVGs) that are interpretable by both AI models and humans, enhancing the reasoning capabilities of AI systems.
- **Efficient Content Creation:** Innovating in the field of conditional visual editing, enabling the manipulation of images and videos based on various parameters for dynamic and real-time applications.

Academic Experience

Human-Readable SVG Generation for Simple Images with Vision Language Models **UIUC**
Assistant Prof. Haohan Wang *June 2023 - Nov. 2023*
PyTorch

- proposed S^2VG^2, the first method combined with a vision language model for SVG generation
- introduced a specialized dataset named SVG-SHAPE, designed for evaluating SVG generation methods and reasoning of LLMs
- demonstrated state-of-the-art performance in SVG reasoning of LLMs and vision metrics

One-shot Controllable Head Avatar with Vertex-feature Transformer **UCI**
Prof. Xiaohui Xie *Apr. 2023 - June 2023*
PyTorch

- proposed CVTHead, a one-shot controllable head avatar framework , which is the first work that performs point-based neural rendering from a monocular face image.
- evaluated our method in comparison to other methods for cross-identity reenactment
- demonstrated state-of-the-art performance on VoxCeleb1 and VoxCeleb2

Trajectory Prediction and Driving Video Caption **AIR, Tsinghua University**
Assistant Prof. Hao Zhao *May 2022 - Sep. 2022*
NumPy, PyTorch

- predicted trajectory on an new interactive motion dataset through AgentFormer and Trajectron++
- trained a novel end-to-end transformer generating descriptions and explanations of driving videos
- demonstrated state-of-the-art performance in driving video captioning

Project Experience

Multimodal Data Synthesis through Entity Detection and Replacement **UCI**
Prof. Xiaohui Xie *June 2023 - Nov. 2023*
PyTorch

- developed an novel method for synthesizing multimodal data through the identification and substitution of entities in text-image pairs , effectively increasing the variety of training data
- demonstrated the effectiveness of the synthesized data in achieved similar performance in tasks like image captioning and visual question answering while only using 75% data of the training set

Professional/Teaching Experience

Lightweight OCR Models Support for OpenCV **Google Summer of Code 2022**
OpenCV *May 2022 - Sep. 2022*
PyTorch, ONNX, C++

- implemented the detection part of PP-OCRv2 model in OpenCV Zoo by ONNX
- implemented high level C++ API of PP-OCRv2 model in OpenCV
- implemented evaluation metrics of text detection (AP, Recall, Precision, Hmean) in OpenCV Zoo

Teaching Assistant for Introduction to Java Programming **Shenzhen, China**
CS102 B, SUSTech *Mar. 2023 - June 2023*
English, Java

- designed and graded a significant portion of the coursework, including assignments and projects.
- developed and managed an online judging platform for evaluating student code submissions

Publications

Tong Zhang , Haoyang Liu , Peiyan Zhang , Yuxuan Cheng and Haohan Wang, "Beyond Pixels: Exploring Human-Readable SVG Generation for Simple Images with Vision Language Models", ***Under Review***, *The IEEE / CVF Computer Vision and Pattern Recognition Conference (CVPR)*, 2024

Haoyu Ma, **Tong Zhang**, Shanlin Sun, Xiangyi Yan, Kun Han and Xiaohui Xie, "CVTHead: One-shot Controllable Head Avatar with Vertex-feature Transformer", *IEEE/CVF Winter Conference on Applications of Computer Vision (WACV)*, 2024

Bu Jin, Xinyu Liu, Yupeng Zheng, Pengfei Li, Hao Zhao,**Tong Zhang**, Yuhang Zheng, Guyue Zhou and Jingjing Liu, "ADAPT: Action-aware Driving Caption Transformer", *IEEE International Conference on Robotics and Automation (ICRA)*, 2023

Awards

- 2nd place of 2022 APAC HPC-AI Competition *Nov. 2022*
- Outstanding Anti-COVID19 Volunteer (SUSTech) *Apr. 2020*

Expert Skills

- Programming Languages: C++, Python, Java
- Libraries/Software: PyTorch, NumPy, Latex