"JACKIE" QIANQI YAN

qyan79@ucsc.edu ◇ 734-707-3373 ◇ Sunnyvale, CA ◇ LinkedIn

EDUCATION

University of California, Santa Cruz — *Santa Cruz, U.S.*
Ph.D. Computer Science and Engineering — *Aug. 2023 - May. 2028 (Expected)*
- Advised by Prof. Xin Eric Wang

University of Michigan — *Ann Arbor, U.S.*
B.S.E. Computer Science, Summa cum laude — *Aug. 2021 - May. 2023*
- Advised by Prof. Joyce Chai, Prof. Stella Yu

Shanghai Jiao Tong University — *Shanghai, China*
B.S.E. Electrical and Computer Engineering — *Sep. 2019 - Aug. 2023*

Technische Universität Berlin — *Berlin, Germany*
Visiting Student — *Jan. 2020 - Feb. 2020*

WORK & TEACHING EXPERIENCE

Initium AI — *Ann Arbor, U.S.*
Machine Learning Software Engineer — *Apr. 2022 - Aug. 2022*
- Incorporate TextRank algorithm and neural network approach to improve ROUGE scores of abstract summarization on DialogSum dataset from 0.44 by 7%.
- Deploy pipeline on TorchServe frontend and transcribe recorded sales conversations from audio to summarized text. Achieve 2-fold decrease in runtime.
- Implement rule-based post-correction algorithm to address incorrect references of people in dialogue summaries.

University of Michigan — *Ann Arbor, U.S.*
WIN. 2023 **Human-Centered Machine Learning (EECS 448)**, Instructional Aide
WIN. 2023 **Intro to Natural Language Processing (EECS 487)**, Grader
WIN. 2022 **Intro to Computer Organization (EECS 370)**, Grader
Shanghai Jiao Tong University — *Shanghai, China*
SU. 2021 **Intro to Circuits (VE 215)**, Teaching Assistant
FA. 2020 **Intro to Computers and Programming (VG 101)**, Teaching Assistant

PROFESSIONAL SERVICE

Nov. 2023 **ICLR 2024**, Reviewer
Oct. 2023 **ICCV 2023**, Reviewer of Workshop on Closing the loop between Vision and Language

RESEARCH EXPERIENCE

Embodied and Responsible Interaction and Communication (ERIC) Lab - University of California, Santa Cruz *Santa Cruz, U.S.*
Advisor: Xin Eric Wang — *Aug. 2023 - Present*

- **Multi-panelVQA: Probing Vision-Language Models on Multi-panel Images**
  - Investigate the capabilities and limitations of Vision Language Models (VLMs) in understanding multi-panel images. Introduce the Multi-panelVQA benchmark, a novel dataset comprising 1,320 image-question-answer triplets, designed to test VLMs in two critical areas: multi-panel image layout understanding ability and multi-panel image content understanding ability. Highlight two key challenges for VLMs in recognizing and interpreting subfigures within multi-panel images: sensitivity to subfigure size and varying effectiveness based on layout complexity.

- **Medical Hallucination Evaluation in Vision-Language Models**
  - Introduce a novel benchmark specifically designed to evaluate the performance of VLMs in generating relevant domain-specific information in addition to employing the CHAIR benchmark to assess object hallucination. Utilizes state-of-the-art (SOTA) general VLM models, as well as those fine-tuned specifically for the medical domain on diverse tasks, including image captioning, medical Visual Question Answering (VQA), and referring expression tasks. Highlight a deeper understanding of the strengths and limitations of VLMs in the critical domain of medical image analysis.

Situated Language and Embodied Dialogue (SLED) Lab - University of Michigan *Ann Arbor, U.S.*
Advisor: Joyce Chai *Jan. 2022 - Feb. 2023*

- **Ground Language and Memory in Robotic Perception and Affordance**
  Presented at the 2022 Microsoft Research Summit (Microsoft Turing Academic Program Workshop).
  - Leverage common sense in large language models (LLM) and incorporate episodic memory & multimodal model to enable embodied agent to outperform state-of-the-art baseline in goal localization and manipulation tasks in AI2-THOR environment. Design a prompting pipeline to query GPT-3 to generate actionable plans based on goal state and environment feedback. Collect 70k image-text pairs of egocentric view and goal state from FILM dataset to fine-tune CLIP model to accurately match goal states with stored frames during inference.
- **Language-Aided Object Detection**
  - Leverage common sense in LLM to improve performance of state-of-the-art object detectors on COCO, PASCAL datasets in a zero-shot manner. Develop a post-correction pipeline for object detectors to infer possible refinement of labels given scene description and spatial relation between bounding boxes.

Stella Yu Group - University of Michigan *Ann Arbor, U.S.*
Advisor: Stella Yu *Sep. 2022 - Mar. 2023*

- **Hierarchical Semantic Segmentation**
  - Build a vision model which conducts image-level recognition and semantic segmentation concurrently by matching hierarchy of image segmentation and language entity at each level of granularity. Extract 50k images from the CC12M dataset which contains hierarchical information in caption using Stanford CoreNLP parser. Introduce extra contrastive loss and fine-tune state-of-the-art model (GroupViT) based on the refined loss.

RELEVANT SKILLS

Programming Languages: Python (PyTorch, TensorFlow, NLTK, Keras), C, C++, C#, SQL, Java, JavaScript
Software & Libraries: Git, Bash, MATLAB, LaTeX, Mathematica, Unity, Unreal Engine, Verilog
Language Skills: English (fluent), Mandarin (fluent), German (beginner)

SELECTED AWARDS AND HONORS

Sep. 2023 **Regents Fellowships**, University of California, Santa Cruz
Sep. 2023 **CSE Deparment Fellowships**, University of California, Santa Cruz
Mar. 2022 **James B. Angell Scholar**, University of Michigan
2022 & 2021 **University Honors**, University of Michigan
Apr. 2022 & Dec. 2021 **Dean's List**, University of Michigan
Apr. 2020 **Undergraduate Excellent Scholarship**, Shanghai Jiao Tong University

SELECTED COMMUNITY SERVICE

2019-2021 **Student Union Sports Department**, President *SJTU*
2019-2021 **UM-SJTU Joint Institute Badminton Club**, President *SJTU*