XIYU LUO

+86 150-5669-3130 | luoxy2022@mail.sustech.edu.cn
+1 (858) 214-5254 | xiyuluoxy@gmail.com

EDUCATION

Southern University of Science and Technology   Shenzhen, China
Bachelor of Engineering in Computer Science and Technology   Sept. 2022 – Jun. 2026
Advisor: Xuan Song
- GPA: 3.81 / 4.00 (Major GPA: 3.91 / 4.00)
- Selected Courses:
  Computer Vision (A+), Natural Language Processing (A), Artificial Intelligence (A),
  Probability and Statistics (A), Data Structures and Algorithm Analysis (A),
  Computer Organization (A+) , Compilers (A), Computer Networks (A).

University of California, San Diego   La Jolla, U.S.
Exchange Program   Sept. 2025 - Dec. 2025
- Courses: Web Mining and Recommender Systems, Machine Learning, Neurobiology of Cognition.

RESEARCH INTERESTS

- *Evaluate and enhance the ability of interaction and alignment of visual and textual modalities, and gain inspirations from cognitive science while doing this.*
- *Unify vision-centered multimodal understanding and generation. Unify representations and frameworks.*
- *Develop multimodal models that support perception, reasoning, memory, and affect-aware understanding in complex environments.*

RESEARCH EXPERIENCE

National University of Singapore, School of Computing   Singapore
Mentor: Colin Tan, [Deep Learning]
Summer Research   Jul. 2024
- Participated in an intensive summer research program focused on Deep Learning.
- Collaborated with a robotics team to develop an intelligent home automation robot capable of plant watering and health monitoring, utilizing YOLO and LSTM, Transformer models.

Southern University of Science and Technology, Department of Computer Science and Engineering   Shenzhen, China
Mentor: Xuan Song, [Human Motion Video Generation: A three-Stage Text-to-Video Pipeline]
Innovative Research   Aug. 2024 – Jan. 2025
- Built a Skeleton-Guided 3-stage text-to-video pipeline. Developed a new metric, HMAD (Human Motion Accuracy Distance), to evaluate human motion quality in generated videos, as an alternative to FID or FVD.
- The HMAD metric provides a comprehensive assessment of both temporal consistency and motion accuracy, with interpretable results through its decomposition of motion and style components.
- Videos generated by our pipeline typically show better HMAD and CLIP scores compared to public models.

University of Oxford, Big Data Institute   Oxford, U.K.
Mentor: Tianyang Zhang, [Translation and Spatial Tracing for Surgical Prediction]
Research Collaboration   Feb. 2025 – Aug. 2025
First author, Accepted to AAAI-26.
- Most existing image-to-image translation methods primarily aim to match the target distribution and often neglect spatial correspondences between the source and translated images. This limitation can lead to structural inconsistencies and hallucinations, undermining the reliability and interpretability of the predictions. These challenges are accentuated in clinical applications by the stringent requirement for anatomical accuracy.
- We present TraceTrans, a novel deformable image translation model designed for post-operative prediction that generates images aligned with the target distribution while explicitly revealing spatial correspondences with the pre-operative input.
- The framework employs an encoder for feature extraction and dual decoders for predicting spatial deformations and synthesizing the translated image. The predicted deformation field imposes spatial constraints on the generated output, ensuring anatomical consistency with the source.

PUBLICATIONS

[1] **TraceTrans: Translation and Spatial Tracing for Surgical Prediction** PDF
Xiyu Luo*, Haodong Li*, Xinxing Cheng, He Zhao, Yang Hu, Xuan Song, Tianyang Zhang.
Accepted to AAAI 2026 (The 40th Annual AAAI Conference on Artificial Intelligence).

PATENTS

[1] *A Method for Medical Image Generation and Postoperative Outcome Prediction Based on Generative Models and Spatial Traceability Techniques.* **Patent Pending**, 2025. *(First Inventor)*

ADDITIONAL INFORMATION

Interests: Photography, Hiking, Surfing, Skiing, Traveling, History, Philosophy, Finance.
Computer Skills: Python (PyTorch), C, C++, Java (SpringBoot), Photoshop, DaVinci Resolve.
Language: Chinese (Native), English (Fluent); Japanese (Conversational); French (Elementary); Spanish (Beginner).