Shun Liu

+1-716-529-8545 | shunliu@buffalo.edu | Homepage
Linkedin | Scholar | Kaggle
Buffalo, New York, United States

Education

- **Shanghai University of Finance and Economics (*Project 211*)** *September 2021 - June 2025*
  B.E., Computer Science Shanghai, China
  - **Relevant Coursework**: Machine Learning, Deep Learning, Social Network Analysis (91), Algorithm Design and Analysis (90), Python (90), Linear Algebra, Discrete Mathematics (99), Advanced Mathematics
  - **Research Interests**: Multi-model Learning, Medical Artificial Intelligence, Automated Machine Learning, Interpretable Computer Vision

Research Experiences

- **A Versatile Framework for Large-Scale Referring Surgical Image Segmentation** *09/2024 - 11/2024*
  Outlet: CVPR'25 (under review) | ***RA at University at Buffalo (SUNY)]*** [GitHub]
  Supervisor: Prof. David Doermann, Dr. Xuan Gong
  - **Task Definition:** Introduced a text-prompt-based segmentation framework for surgical images under text supervision, composed of various fine-grained attributes, and instruction complexities.
  - **Benchmark Development:** Built a large-scale dataset (66K images, 242K masks, 1M instructions) with fine-grained annotations (e.g., color, size, location) from scratch and construct a holistic categorization over 76 classes to map with anatomical structures.
  - **Generalization:** Showcased strong cross-dataset generalization and open-vocabulary segmentation capabilities over competitive baselines like GroundedSAM.

- **A Large-scale Vision-Language Dataset for Endoscopic Surgery Understanding** *04/2024 - 08/2024*
  Outlet: ICLR'25 (under review) | ***RA at University at Buffalo (SUNY)]*** [GitHub]
  Supervisor: Dr. Xuan Gong
  - **Dataset Creation:** Collected surgery videos from open sources (YouTube, MedTube) and annotated video frames with clinical Q&A data with the largest scales and the professional questions are aligned with endoscopy experts.
  - **Experimental Performance:** Finetuned LLaVA on the new dataset, achieving SOTA performance on biomedical tasks like zero-shot classification and text-image retrieval, such as Kvasir, Hyper-Kvasir, GastroVision, NBI-Inframes.

- **DAG-Driven Protein Sequence Representation and Function Prediction** *Jul 2023 - Dec 2023*
  Outlet: Gold Medal (Top 15/1625 teams) | *Challenge Announcement* [GitHub]
  Solo Team
  - **Overview.** Developed directed acyclic graphs (DAG) for Gene Ontology (GO) to represent biological processes, cellular components, and chemical functions of proteins, uncovering semantic associations.
  - **Proposed Approach.** Mapped DAG subsets to specific protein functions, integrating ProFun, QuickGO, and SPROF codes to enhance GO graph reconstruction over sequence-based methods.

- **Value System and Potential Group-Dependent Bias in LLMs** *10/2023 - 01/2024*
  Outlet: Conference Manuscipt | ***RA at Dartmouth College*** [GitHub]
  Supervisor: Prof. Sorough Vosoughi
  - **LLM Stereotyping Biases**. This study explores biases in large language models (LLMs) across five key attributes: fairness, reliability, robustness, privacy, and interpretability, examining how LLMs' attitudes toward different groups are influenced by human values.
  - **Qualitative Analysis**. Qualitative experiments on multilingual LLMs test theoretical assumptions about group-dependent values and conduct ablation studies on RLHF, raising promising research questions.

- **A Real-time Yet Memory-Efficient Medical Imagery Detection Model** *09/2023 - 11/2023*
  Outlet: Preprint at arxiv [GitHub]
  Supervisor: Prof. Teok Teik Toe
  - **Real-Time Multi-Object Detector**. Based on the insufficiency of YOLOv8 algorithm in multi-scale target detection, a dynamic adaptive detection head is proposed to solve the variability of object sizes.
  - **Class Imbalance**. Combined with the DFL loss, the phenomenon of uneven-distributed medical images in BCCD dataset is handled, results have proven the effectiveness (mAP@50 above 0.90) and its extensive performances on remote-sensing fields, evaluated by MAR20 benchmark.

PUBLICATIONS

C=CONFERENCE, J=JOURNAL, P=PREPRINT, S=IN SUBMISSION, W=WORKING PAPER

[W.1] **Shun Liu**, Shuting He. **R2-HOI: Exploring the Transferability of Referring Video Object Segmentation (RVOS) Models for Open-Set Human-Object Interaction Detection**. ***ICML'25***.

[S.1] **Shun Liu**, Nan Xi, Yang Liu, Tianyu Luan, Chenwei Wu, Yunjie Tian, Xuan Gong, David Doermann. **A Versatile Framework for Referring Segmentation with Large-Scale Surgical Endoscopy Images**. ***CVPR'25***.

[S.2] Xuan Gong, Balu Harshavardan Koduru, Yuanhao Zhai, **Shun Liu**, Nan Xi, Xi Tang, Yuan Zhang, Tenzin Lhakpa, Yunjie Tian, Yuxuan Sun, Tianyu Luan, Ziqing Xue, Junsong Yuan, David Doermann. **EndoAssistant: A Large-scale Vision-Language Dataset for Endoscopic Surgery Understanding from Open-Source Videos**. ***ICLR'25***.

[C.1] Nguyen Minh Thao Phan*, Cong-Tinh Dao*, Chenwei Wu, Jian-Zhe Wang, **Shun Liu**, Jun-En Ding, David Restrepo, Feng Liu, Fang-Ming Hung, Wei-Chih Peng. **MedFuse: Multimodel EHR Data Fusion with Masked Lab-Test Modeling and Large Language Models**. ***CIKM'24 (Short Research Paper Track, accept rate 27%)***.

[S.3] Weicheng Ma, Ethan Gearey, James Quirk, **Shun Liu**, Lili Wang, Soroush Vosoughi. **Exploring Language and Model-Specific Biases in LLM Stereotyping Behaviors**. ***EMNLP'24***.

[P.1] **Shun Liu**, Jianan Zhang, Ruocheng Song, Teik Toe Teoh. **ADA-YOLO: Dynamic Fusion of YOLOv8 and Adaptive Heads for Precise Image Detection and Diagnosis**. Preprint available at arxiv.

[P.2] **Shun Liu**. **Model-Agnostic Interpretation Framework in Machine Learning: A Comparative Study in NBA Statistics**. Preprint available at arxiv.

INDUSTRY EXPERIENCES

- **Cardinal Operations [🌐]** *03/2024 - 06/2024*
  Research Intern (Group of Large Language Models Technologies) Shanghai, China
  - **LLM-Driven Forecasting Systems:** Designed a semi-supervised feature engineering pipeline for heterogeneous high-dimensional tabular data with LLMs for forecasting tasks across retail, manufacturing, and energy industries.
  - **Performance Impact:** Achieved best WMAPE scores in time-series forecasting tasks across multiple industrial datasets, including the forecasting over sales, house prices, manufacturing parameters.
- **Zhejiang Lab [🌐]** *08/2023 - 01/2024*
  Research Intern (Institute of Artificial Intelligence) Hangzhou, China
  Supervisor: Dr.Hongsheng Wang & Prof. Shengyu Zhang
  - **Academic Writing:** Contributing to the patent and academic drafting, studied the fundamental knowledge of human joint rotation distribution model (exemplar: Kinetic Tree).
  - **Research Training:** Conducted cutting-edge researches on diffusion-guided human mesh recovery and flow-based motion reconstruction within private rehabilitation data sampled from local hospitals.

SKILLS

- **Programming Languages:** Python (Proficient), C++ (Intermediate), MATLAB (Beginner)
- **Tools:** Git, LaTeX (Overleaf), Linux Shell (Bash/Zsh)
- **Languages:** English (IELTS: 6.5/9, Duolingo: 130/160)

CHALLENGES AND AWARDS

- **CAFA 5 Protein Function Prediction** *2023*
  Ranked 15/1625 (Top 0.9%), Gold Medal (Solo) [🌐]
  - Decompose protein structures using graph representation and Gene Ontology (GO) domain knowledge, then make accurate and robust function prediction.
  - Outperformed over 99% of teams in designing GO graph representation for robust protein function prediction.
- **Large Language Models for Science Exams** *2023*
  Ranked 50/2664 (Top 1.9%), Silver Medal [🌐]
  - Finetune small LLMs using RAG techniques to better answer STEM-related queries.
- **Linking Writing Processes to Writing Quality** *2023*
  Ranked 144/1876 (Top 7.6%), Silver Medal [🌐]
  - Model and connect personal typing habits with essay quality assessment.