DSA Research Experiences for Undergraduates

Research Project

Section1: Faculty Information

Full Name	Weikai YANG	Tel	13051670559
Thrust/Hub	DSA/INFH	Office	W1-316
Email	weikaiyang@hkust-gz.edu.cn		

Section2: Research Project Proposal

Project Title	Pictorial Chart Understanding
Project Description (max 800 words)	1. Background Pictorial charts are favored for their memorability and visual appeal, offering a more engaging alternative to basic charts. However, existing visual language models or Chart QA models fail to accurately understand pictorial charts and answer related questions due to the substantial domain gap difference between the basic charts and fancy charts. While finetuning with pictorial charts is a promising way, it is costly to collect sufficient pictorial charts and annotate them for training. To address this need, this project aims to synthesis pictorial charts using generative models and use them to enhance the model capability in understanding pictorial charts in real world. 2. Challenges (a) **Data Scarcity and Annotation Costs:** Collecting and annotating large-scale real-world pictorial charts is labor-intensive due to their niche use cases and copyright restrictions. (b) **Domain Gap Complexity:** Pictorial charts (e.g., infographics, illustrated diagrams) exhibit highly stylized visual features (e.g., metaphors, icons, decorative elements) that differ radically from basic charts, causing catastrophic failures in existing Chart QA models. 3. Objectives (a) Develop a generative framework to synthesize high-fidelity, diverse pictorial charts with controllable attributes (e.g., iconography, layout, thematic styles). (b) Design a domain adaptation pipeline to bridge the gap between synthetic and real-world pictorial charts, enhancing model generalizability. (c) Achieve state-of-the-art performance on pictorial Chart QA tasks using synthetic data-augmented training. 4. Methodology (a) Controllable Pictorial Chart Synthesis. There are two ways to achieve this: 1) Leverage diffusion models or GANs to generate pictorial charts with fine-grained control over, and 2) Use visual motifs (e.g., icons,

	metaphors) and rule-based rendering techniques to generate them. After generating pictorial charts, we will also generate associated question-answer pairs for training and testing. (b) Generate stylistic variations (e.g., flat design, 3D realism) using LoRA adapters. (c) Train a vision-language model (VLM) on both basic charts and pictorial charts, and apply domain-invariant contrastive learning to align synthetic and real chart embeddings. 5. Expected Outcomes (a) An open-source generative toolkit for synthesizing stylized pictorial charts with customizable attributes (b) A benchmark dataset of 1,000+ real and synthetic pictorial charts with annotated QA pairs. (c) A fine-tuned vision-language model achieving >20% improvement in QA accuracy on pictorial charts over baseline models. (d) Guidelines for optimizing synthetic data realism/diversity trade-offs in low-resource visual domains.
Proposed Research Duration	Start Date: ___MAR_____ / ___15_____ / ____2025____ End Date: ___SEPT_____ / ___15_____ / ____2025____
Student/Researcher Duties	- Actively participate in discussions and provide regular feedback - Conduct a literature review on existing Chart QA method and pictorial chart generation method - Implement the toolkit for pictorial charts generation - Train the model on the constructed dataset, then design and conduct rigorous experiments to evaluate the performance
Technical Skills Required	☑ Python ☑ Machine Learning ☐ Big Data ☐ R ☑ Deep Learning ☐ SQL ☐ C/C++ ☐ Other: _____
Preferred Student/Researcher Background	Python, and the experience of running generative models. But it is not required as I will provide necessary guidance to help you.
Maximum Number of Students/Researchers	☐ 1 ☑ 2

Section3: Pre-Application Research Exposure Meeting

Faculty members are encouraged to schedule a Research Exposure Meeting to introduce students to their projects.

Preferred Date	Discussed via email
Preferred Time	Discussed via email
Meeting Mode	☑ In-Person ☐ Online
Venue (if in-person)	W1-316
Meeting Link (if online)	