XINYUAN WANG

3869 Miramar St., La Jolla, USA

📞 +86 158 4837 4206   ✉ xiw136@ucsd.edu   ⭘ https://xinyuanwangcs.github.io/

Education

University of California San Diego (UCSD)   **Sep. 2022 – June 2024 (expected)**
Master of Science in Computer Science   *San Diego, California, USA*
- GPA: 4.00/4.00
- Research Mentors: Prof. **Zhiting Hu**, Prof. **Zhuowen Tu**

Central South University (CSU)   **Sep. 2018 – June 2022**
Bachelor of Science in Computer Science and Technology   *Changsha, Hunan, China*
- GPA: 90.97/100 (3.88/4.00)
- Research Mentor: Prof. **Ying Zhao**, Sub-Dean of the Department of Computer Science and Technology
- Awarded 2022 Outstanding Graduates of Central South University (Jun. 2022)
- Awarded the School of Computer Science and Engineering Annual Scholarship (2018 - 2021)

Research Interests

- **Large Language Models (LLMs) with World Models**: Augmenting LLMs with interfaces to world models and mechanisms to understand and process world information, thereby enabling them to address real-world challenges more effectively. (LLM-Reasoners)
- **Foundation Model Inference Enhancement**: Foundation models demonstrate remarkable capabilities across a range of tasks. However, due to limitations in how we interact with them, their inference abilities are not fully exploited. One promising direction is to explore expert-level prompt optimization, bridging the gap between the model's inherent distribution and the specific distribution of tasks (PromptAgent)
- **Conceptual Enhancement of Generative Models**: While large generative models can produce data with high fidelity and diversity, but they can only learn conceptual alignment implicitly. This can result in conceptual errors and incorrect heuristics. By focusing on conceptual enhancement, we can improve the reliability and interpretability of these models.

Research Experiences

PromptAgent: Automatic LLM Prompt Optimization Framework   **May 2023 – Present**
*Mentored by **Zhiting Hu**, Assistant Professor at HDSI and CSE Department*   *University of California, San Diego*
- **Motivation**: Existing prompt optimization methods tend to overlook the depth of domain knowledge and struggle to efficiently explore the vast space of expert-level prompts using simple sampling or searching methods. PromptAgent formulates prompt optimization as a strategic planning problem and employs a principled planning algorithm to navigate the expert-level prompt space efficiently and effectively.
- **Methodology**: Introduced a novel automatic prompt optmization agent that integrates Monte Carlo Tree Search (MCTS), a principle planning algorithm, to explore the prompt space. Additionally, by employing a trial-and-error approach, the domain knowledge in the data is incorporated into the optimzed prompt during optimizing.
- **Results**: PromptAgent has been rigorously experimented on various tasks, covering BIG-Bench Hard, domain-specific, and general NLP tasks, showing it outperforms strong human prompting methods, such as Chain-of-Thought, as well as recent prompt optimization baselines. Analyses highlight its ability to craft expert-level, detailed, and domain-insightful prompts with superior performance and efficiency than other baselines. Moreover, the optimized prompts demonstrate better generalizability across models compared to those written by humans or other baselines.
- **Preprint**: PromptAgent: Strategic Planning with Language Models Enables Expert-level Prompt Optimization
  Under Review of ICLR 2024 [Arxiv]

LLM Reasoner Library   **August 2023 – Present**
*Research assistant to **Zhiting Hu**, Assistant Professor at HDSI and CSE Department*   *University of California, San Diego*
- **Motivation**: LLM Reasoners is a library designed to enhance the reasoning capabilities of Large Language Models (LLMs) using advanced algorithms. It treats multi-step reasoning as a planning task, searching for the optimal reasoning chain. This approach aims to strike the right balance between exploration and exploitation, guided by concepts like the 'World Model' and 'Reward'.
- **Methodology**: The library is based on Reasoning via Planning. It offers the most up-to-date search algorithms for reasoning with LLMs, such as RAP-MCTS, Tree-of-Thoughts, Guided Decoding, and more. Additionally, the framework is compatible with various LLM platforms, including Huggingface transformers, OpenAI API, and others. I am responsible for: 1. Managed experiments on Math datasets using various reasoning methods, including Least-to-Most Prompting, CoT, and RAP-MCTS. 2. Developed Tree-of-Thought searching algorithm.

- **Results**: This library integrates cutting-edge LLM reasoning algorithms, the latest LLM models, and intuitive visualization. It also offers accessible interfaces for all incorporated algorithms and models. Notably, among the recent LLM reasoning algorithms, our RAP-MCTS algorithm shows superior performance over other searching or prompting methods on Math, Logical, and Embodied tasks.
- **Library**: LLM Reasoners **GitHub Stars**: **521**

Refining Diffusion Model Loss with End-to-End Information   July 2023 − Present

*Mentored by **Zhuowen Tu**, Professor at Cognitive Science and CSE Department*   *University of California, San Diego*

- **Motivation**: Diffusion models can generate photorealistic images, but they often make many semantic mistakes, such as twisted fingers and ears. This is because Diffusion models are trained with per-pixel loss (MSE) in an auto-regressive manner, which lacks end-to-end conceptual information. Incorporating end-to-end information into the training loss of diffusion models is essential to correct the conceptual errors and enhance performance of diffusion models.
- **Methodology**: A Generative Adversarial Model (GAN) is concurrently trained with the Diffusion Model, utilizing end-to-end synthesized data. The discriminator of GAN injects conceptual guidance into Diffusion models' training process. This approach allows the original per-pixel loss to be augmented with comprehensive set information.
- **Results**: Experiments show that the unconditional transformer-based Diffusion model, DiT, can generate images of higher fidelity (measured by FID) when trained with end-to-end conceptual guidance on CelebA-HQ Dataset.

Interpretable Object Detection via Deep Learning   Dec. 2020 − May 2022.

*Mentored by **Ying Zhao**, Sub-Dean of the Department of Computer Science and Technology*   *Central South University*

- **Motivation**: CNNs are widely used as black boxes, while the interpretability of their features and feature importance is less explored. This project aimed to improve the interpretability of CNN by visualization and provide guidance for feature transferring of high-value features in CNN.
- **Methodology**: Proposed the Average Image Analysis method to evaluate the network's low-frequency information. It measures the input's influence on neurons by calculating the cosine distance between neuron's guided-backpropagation visualizations and the dataset's average image.
- **Results**: This approach improves the clarity of gradient-based CNN visualizations, deepens insights into semantic information at different network depths, and provides guidance for subsequent feature transfer. Experiments reveal that pruning 40% of high-importance features dramatically reduces performance, while eliminating 40% of low-importance features has negligible impact on the model's performance.
- **Software Copyright**: A Visual Analysis of Internal Feature Importance and Feature Transfer Method in Convolutional Neural Networks, 2021-06-25
- **Graduation Thesis**: The Research on The Interpretability Method of Deep Neural Network Based on Average Image

Publications and Software Copyrights

Publications (* equal contribution)**:**

- **PromptAgent: Strategic Planning with Language Models Enables Expert-level Prompt Optimization**
  Xinyuan Wang*, Chenxi Li*, Zhen Wang*, Fan Bai, Haotian Luo, Jiayou Zhang, Nebojsa Jojic, Eric Xing, Zhiting Hu
  Under Review of ICLR 2024 [Arxiv]
- **Reduce the medical burden: An automatic medical triage system using text classification BERT based on Transformer structure**
  Xinyuan Wang*, Make Tao*, Runpu Wang*, Likui Zhang*
  Proceedings of ICBASE 2021 [Paper]
- **A Fast Method for Detecting Minority Structures in a Graph**
  Fangfang Zhou, Qi'an Chen, Yunlong Cui, Xinyuan Wang, Hongxu Ma, Ying Zhao, Xiaoli Li
  Proceedings of VINCI '20 [Paper]

Software Copyrights:

- A Visual Analysis of Internal Feature Importance and Feature Transfer Method in Convolutional Neural Networks
  2021107101643, 2021-06-25
- Interactive Family-building Software of Malicious Webshell Files V1.0
  2022SR0850594, 2022-06-27
- Application Layer Malicious Network Traffic Detection Software Based on the Cascaded Deep Neural Network
  2021SR0330731, 2021-01-02

Awards

- 2022 Outstanding Graduates of Central South University   *Jun. 2022*
- Guangyun Technology Scholarship (4 out of 196 students)   *Dec. 2021*
- School of Computer Science and Engineering Annual Scholarship   *2018 - 2021*

Activities

- **Reviewer**: ICLR 2024, AAAI 2024, NLPCC2023
- **Blogs**: CSDN