

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 a world model formulation to enable principled decision-making, planning, and simulation. Enhancing the LLM's abilities in reasoning, planning, and interacting with the world. ([LLM Reasoners](#))
- **Foundation Model Prompting:** Employing interpretable prompting to bridge the domain gap between user objectives and the outputs of foundation models. Effectively boosting the performance of foundation models on complex tasks through efficient and effective prompting. ([PromptAgent](#))
- **Semantic Enhancement and Control of Generative Models:** Generative models, such as text-to-image models, sometimes exhibit semantic inconsistencies and challenges in control. My aim is to integrate semantic information into the models during training or inference to enhance their semantic fidelity, reliability, and controllability.

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

- We aim to bridge the divide between LLMs and domain-specific tasks for users. To achieve this, we use an LLM Agent to efficiently navigate the prompt space and effectively extract crucial data insights. We proposed PromptAgent, a novel prompt optimization agent, by formulating prompt optimization as a strategic planning problem and employing a principled planning algorithm to traverse the expert-level prompt space efficiently and effectively.
- PromptAgent is rigorously experimented on various tasks showing it outperforms strong human prompting methods and recent prompt optimization baselines. Analyses highlight its ability to craft expert-level, detailed, and domain-insightful prompts with superior performance, searching efficiency, and generalizability.
- **Preprint:** PromptAgent: Strategic Planning with Language Models Enables Expert-level Prompt Optimization
Under Review of [ICLR 2024](#) [[Arxiv](#)]

Instruction Tuning with PromptAgent

Oct. 2023 – Present

Mentored by [Zhiting Hu](#), Assistant Professor at HDSI and CSE Department

University of California, San Diego

- Instruction tuning (IT) is essential for LLMs to align human expectations and LLMs' outputs, but it is less effective in smaller-scale models. We aim to improve the instruction understanding and following ability of relatively small models, such as Alpaca 7B. We leverage PromptAgent to generate instructions that encompass quantity, diversity, and task-specific knowledge across various tasks. By tuning with these intricate prompts imbued with expert-level domain insight, we aim to enable the model to better adhere to complex instructions and potentially foster advanced instruction-following capabilities like reasoning. (This is an on-going following work of PromptAgent.)

LLM Reasoners (500+ Stars)

August 2023 – Present

Research assistant to [Zhiting Hu](#), Assistant Professor at HDSI and CSE Department

University of California, San Diego

- 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. It offers the most up-to-date search algorithms for reasoning with LLMs, such as RAP-MCTS, Tree-of-Thoughts, and is compatible with various LLM platforms, including Huggingface transformers, OpenAI API, and others. I conducted experiments on Math datasets, such as AQuA, using various reasoning methods, including [Least-to-Most Prompting](#), [CoT](#), and [RAP-MCTS](#). I also developed [Tree-of-Thought](#) searching algorithm, which is one of the base searching algorithms in the library.

- 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

- While diffusion models can produce photorealistic images, they frequently commit semantic errors, such as producing twisted fingers. One reason is that their training solely on per-pixel loss (MSE), which lacks comprehensive end-to-end semantic information. We employ a discriminator iteratively trained on end-to-end synthesis data to provide adversarial guidance during the training of diffusion model. This novel training objective strikes a balance between image fidelity, achieved through per-pixel loss, and semantic accuracy, ensured by end-to-end discriminative loss.
- Experiments show that the unconditional Diffusion model, [DiT](#), can generate images of higher fidelity (measured by FID) when trained with end-to-end discriminative guidance on CelebA-HQ Dataset. This project is still on-going.

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

- CNNs are widely used as black boxes, while the interpretability of their features is less explored. We aimed to improve the interpretability of CNN's features by quantifying the feature importance. I proposed the Average Image Analysis method to evaluate the network's low-frequency information by calculating the distance between neuron's feature visualization and the dataset's average images. I also conducted CNN pruning experiments to validate the effectiveness of the metric.
- This approach quantifies semantic information and feature importance at different network depths, and provides guidance for subsequent feature transferring. Experiments show that pruning high-importance features dramatically reduces performance, while eliminating 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
Reviewed by the University Committee [\[Paper\]](#)

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

Skills & Activities

- **Reviewer:** [ICLR 2024](#), [AAAI 2024](#), [NLPC2023](#)
- **Blogs:** [CSDN](#)
- **Selected Courses:** Statistical Natural Language Processing: A+, Unsupervised Learning: A+, Deep Learning for 3D Data: A+