Zhe Li

Email: keycharon0122@gmail.com | Phone: +86 19974863980
Address: Huazhong University of Science and Technology, Hongshan District, Wuhan City, Hubei Province, China 430074

EDUCATION

Huazhong University of Science and Technology **09**/2022 - present

Master's degree in computer science and technology; Average Score: 87.0/100

Modules Include: deep learning, machine learning, reinforcement learning, natural language processing, computer vision, high-performance storage, matrix theory.

Hunan University 09/2017 - 07/2021

Bachelor's degree in biotechnology; Average Score: 79.5/100

Modules Include: advanced mathematics, linear algebra, probability theory, animal biology, plant biology, neurobiology, biochemistry, microbiology, genetic biology.

MANUSCRIPTS

Li, Z., *et al.* (2024). *DCS-Net: Pioneering Leakage-Free Point Cloud Pretraining Framework with Global Insights* [Manuscript submitted for publication]. **First author.**

- Existing point cloud models suffer from information leakage due to the pre-sampling of centroids, which leads to the simplistic task of proxying the model. These methods focus primarily on local feature reconstruction, limiting their ability to capture global patterns within the point cloud. In this paper, we argue that the reduced difficulty of the proxy task hinders the ability of the model to learn expressive representations. To address these limitations, we introduce a novel solution known as a differentiable central sampling network (DCS-Net). It solves the information leakage problem by combining global and local feature reconstruction into a significant agent task, thus enabling the simultaneous learning of global and local patterns within a point cloud.

Li, Z., *et al.* (2024). *General point model with autoencoding and autoregressive* [CVPR'24 Accepted]. **First author.**

- Inspired by the Generalised Language Model, we present the Generalised Point Model (GPM), which seamlessly integrates autocoding and autoregressive tasks in a point cloud converter. The model is versatile, allowing fine-tuning of downstream point cloud representation tasks as well as unconditional and conditional generation tasks. GPM enhances mask prediction in autocoding with various forms of mask-filling tasks, thereby improving the performance of point cloud understanding. In addition, GPM demonstrates highly competitive results in unconditional point cloud generation tasks and even shows the potential of conditional generation tasks by modifying the input conditional information. Compared to models such as Point-BERT, MaskPoint and PointMAE, our GPM achieves excellent performance in point cloud understanding tasks.

Li, Z., *et al.* (2024). *MLIP: Enhancing Medical Visual Representation with Divergence Encoder and Knowledge-guided Contrastive Learning* [CVPR'24 Accepted]. **First author.**

- Existing studies have neglected the multi-granular nature of medical visual representations and lacked suitable contrast learning techniques to improve the generalisation ability of models at different granularities, leading to under-utilisation of image text information. To address this problem, we propose MLIP, a novel framework that utilises domain-specific medical knowledge as a guiding signal to integrate linguistic information into the visual domain through image-text contrast learning. Our model consists of global contrast learning using our designed divergence encoder, local token-knowledge-patch alignment contrast learning, and knowledge-guided category-level contrast learning using expert knowledge.

Li, Z., *et al.* (2023). Enhancing Sentence Representation with Visually-supervised Multimodal Pre-training. In *Proceedings of the 31st ACM International Conference on Multimedia* (pp. 5686-5695). **First author.**

- Most pre-trained models currently use transformer-based encoders with a single modality and are primarily designed

for specific tasks such as natural language inference and question-answering. Unfortunately, this approach neglects the complementary information provided by multimodal data, which can enhance the effectiveness of sentence representation. To address this issue, we propose a Visually-supervised Pre-trained Multimodal Model (ViP) for sentence representation. Our model leverages diverse label-free multimodal proxy tasks to embed visual information into language, facilitating effective modality alignment and complementarity exploration. Additionally, our model utilises a novel approach to distinguish highly similar negative and positive samples. We conduct comprehensive downstream experiments on natural language understanding and sentiment classification, demonstrating that ViP outperforms both existing unimodal and multimodal pre-trained models.

PROFESSIONAL EXPERIENCES

Westlake University, China 07/2023 – 11/2023

Part-time Internship

- **Carried out the academic exchange in Stan Z.Li's group and completed a paper on 3D point cloud modelling, which has been accepted by CVPR'24.**

Ningbo Boden AI Technology Co., Ltd. 07/2020 – 09/2020

Part-time Algorithm Internship

- **Utilised semantic segmentation and target detection techniques to divide the drivable area of the road and detect some obstacles on the roadside as well as traffic signs. Employed state-of-the-art computer vision methodologies to identify and detect various obstacles located on the roadside, including vehicles, pedestrians, and roadside barriers.**

Laiye Technology (Beijing) Co., Ltd. Hunan Branch 06/2018 – 08/2018

Part-time Algorithm Internship

- **Identified key triggers within the WeChat environment and implemented automated functionalities for UIBOT, enabling seamless operations within the WeChat platform.**

PROJECT EXPERIENCES

Modelling Theory and System Design for Safety Critical Information Physical Systems

Supported by the National Natural Science Foundation of China (NSFC)

- **Conducted unified modelling of discrete computational and continuous physical systems in ACPS.**

Research on Key Technologies of Artificial Intelligence Basic Models

- Responsible for implementing new basic models of self-supervision and weak supervision for cross-language and cross-modality.

EXTRACURRICULAR ACTIVITIES

- Year Leader of Biology College of Hunan University
- Captain and Coach of the Basketball Team, College of Biology, Hunan University

AWARDS & HONOURS

- Top Ten Outstanding Papers Award 01/2024
- National Scholarship 10/2023
- Outstanding Graduate Student Honour of Huazhong University of Science and Technology 09/2023
- Outstanding Athlete of Hunan University 06/2019
- Champion of Hunan University Basketball Competition 11/2018, 11/2020

ADDITIONAL INFORMATION

- **Languages:** Chinese (Native); English (Proficient)

- **Skills:** Leadership, Management and Organisational Skills; Excellent Communication Skills
- **Software:** Microsoft Office (PPT, Excel, Word)
- **Hobby:** Basketball
- **Other Information:** During my undergraduate studies, I set up a studio and participated in competitions as well as engaged with project development with studio members.