CAN CUI

Google Scholar ◇ Github ◇ Homepage

Phone: (+86) 153-1611-6391 ◇ Email: `cuican@westlake.edu.cn`

EDUCATION

Shanghai Jiao Tong University, Shanghai, China   Sep. 2020 - Jun. 2023
Master of Engineering in Control Engineering
Thesis: Multi-Label Remote Sensing Image Retrieval and Its Application
Supervisors: *Prof. Tao Fang*
GPA: 3.56/4.0

Donghua University, Shanghai, China   Sep. 2016 - Jun. 2020
Bachelor of Engineering in Automation
GPA: 3.56/5.0 (Top 20%)

EMPLOYMENT

Machine Intelligence Lab (MiLAB), Westlake University   Aug. 2023 - Present
Research Assistant
Focuses on tasks related to multi-modality, generative models, and embodied AI.

PUBLICATIONS

[1] **Can Cui**, S. Huang, and et al., "ProFD: Prompt-Guided Feature Disentangling for Occluded Person Re-Identification," in *ACM Multimedia (ACM MM)*, 2024, [Link].
[2] **Can Cui**, H. Huo, and T. Fang, "Deep Hashing with Multi-Central Ranking Loss for Multi-Label Image Retrieval," *IEEE Signal Processing Letters (SPL)*, 2023, [Link].

RESEARCH EXPERIENCE

A Faster and Stronger Diffusion-based VLA Model   Jun. 2024 - Present
Supervisors: Prof. Donglin Wang   Westlake University

Existing diffusion-based policy models perform well but have slow inference speeds, which limits their use in real-world scenarios. Our goal is to propose **a faster and stronger diffusion-based V**ision-**L**anguage-**A**ction **(VLA)** model.

· Designing an asynchronous hierarchical VLA model, where high-level large model is responsible for better understanding instructions, while low-level small model is responsible for faster and more accurate action prediction.
· Conducting experiments to validate the effectiveness of consistency distillation in distilling the 3D Diffuser Actor, enabling one-step trajectory generation.
· Exploring various methods (LISA, Average Pooling) to integrate the reasoning and planning capabilities of the Large Vision-Language model, LLaVa-Next-8B.
· Planning to submit to CVPR 2025.

ProFD: Prompt-Guided Feature Disentangling for Occluded Person ReID[1]   Nov. 2023 - May 2024
Supervisors: Dr. Siteng Huang and Prof. Donglin Wang   Westlake University

ProFD utilizes CLIP to **incorporate textual priors** in Occluded Person ReID, a pure vision task, addressing the two critical problems, missing part information and noisy pseudo-label, in previous pure vision methods.

· Developed a novel framework, ProFD, to address occlusion challenges in person Re-Identification (ReID) by leveraging prompt-guided feature disentangling.
· Introduced part-specific prompts and a hybrid-attention decoder, combining spatial-aware and semantic-aware attention to generate well-aligned part features, enhancing model performance in occluded scenarios.

- Employed a self-distillation strategy to retain the pre-trained knowledge of CLIP, avoiding catastrophic forgetting and improving generalization.
- Conducted extensive evaluations to validate the effectiveness of the proposed methods, achieving state-of-the-art results on multiple public occluded person ReID datasets.

Deep Hashing with MCR Loss for Multi-Label Image Retrieval [2]   Oct. 2022 - Feb. 2023
Supervisors: Prof. Hong Huo and Prof. Tao Fang   Shanghai Jiao Tong University

To tackle position conflicts and multi-level similarity in proxy-based methods for multi-label image retrieval, we introduce MCR Loss.

- Developed a novel deep hashing model with a Multi-Central Ranking Loss (MCR Loss) designed to enhance multi-label image retrieval by preserving complex semantic correlations and reducing quantization error.
- Proposed a multi-central similarity loss using learnable hash centers to optimize the metric space, addressing embedding conflicts caused by proxy-based supervision.
- Introduced a smooth-WAP ranking loss, an approximation of the weighted Average Precision, to preserve multi-level similarities in the metric space.
- Conducted extensive experiments to validate the effectiveness of the proposed loss functions.
- Achieved state-of-the-art performance on two widely-used benchmark datasets (NUS-WIDE and Flickr25K), outperforming existing hashing methods in various ranking evaluation metrics.

PROJECTS

National Key R&D Program of China: Remote Sensing Image Retrieval Platform   Fall Semester, 2021
Student Leader   Shanghai Jiao Tong University

The platform is able to manage large amount of remote sensing images by various retrieval methods. My contributions include:

- Built the client side dynamic web pages with Vue, Node.js and Javascript.
- Created the server side endpoints using Flask and Python.
- Integrated multiple types of image retrieval algorithms in the platform.
- enabled various retrieval methods for remote sensing image, including cross-modal retrieval, cross-source retrieval, multi-label retrieval, and re-ranking of results.

ACHIEVEMENTS

The Second Prize Scholarship, awarded by Shanghai Jiao Tong University	Fall 2020 - Summer 2023
Outstanding Graduates , awarded by Donghua University	Summer 2020
Academic Excellence Prize, awarded by Donghua University	Fall 2019
Outstanding Engineer Scholarship, awarded by Donghua University	Fall 2018
National Second Prize, the 13th "NXP Cup" National Smart Car Design Competition	Fall 2018

SKILLSET

Programming Languages: Python, C/C++, JavaScript.
Tools & Platforms: Pytorch, Hugging Face, Docker, Matlab, Git, Vue, Node.js, flask, Linux.
Machine Learning: Imitation Learning, Generative Model (Diffusion Model, GAN), Large Language Model, CLIP, Parameter-Efficient Fine-Tuning (LoRA, Prompt Tuning, Adapter), Model Acceleration (Mamba, Mixture-Of-Depth, Consistency Model).
Languages: Chinese(native), English (IELTS: 6.5, with L:7.0, R:6.5, W:6.5, S:6.0).