

CAN CUI

Google Scholar \diamond Github \diamond Homepage

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

EDUCATION

Shanghai Jiao Tong University, Shanghai, China

Sep. 2020 - Jun. 2023

Master of Engineering in Control Engineering

Advisors: Prof. Tao Fang and Prof. Hong Huo

GPA: 3.56/4.0

Donghua University, Shanghai, China

Sep. 2016 - Jun. 2020

Bachelor of Engineering in Automation

GPA: 3.6/5.0(top 20%)

EMPLOYMENT

Machine Intelligence Lab (MiLAB), Westlake University

Aug. 2023 - Present

Research Assistant

Focuses on tasks related to computer vision, 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

Acceleration of Inference in Diffusion Policy

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 some scenarios. Our goal is to propose a **faster and better diffusion-based Vision-Language-Action (VLA)** model.

- Designed an asynchronous hierarchical VLA model, where the high-level large model is responsible for better understanding instructions, while the low-level small model is responsible for faster and more accurate action execution.
- Conducted experiments to validate the effectiveness of consistency distillation in distilling the 3D Diffuser Actor, enabling one-step action 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 modality priors** in Occluded Person ReID, a pure vision task, addressing the two critical problems of 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.
- Achieved state-of-the-art results on multiple ReID datasets, including Market1501, DukeMTMC-ReID, Occluded-Duke, Occluded-ReID, and P-DukeMTMC, demonstrating significant improvements in mAP and Rank-1 accuracy.
- Conducted extensive evaluations to validate the effectiveness of the proposed methods, setting new benchmarks in occluded person ReID tasks.

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.
- 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.
- Conducted extensive experiments, including ablation studies, to validate the effectiveness of the proposed loss functions.

ACHIEVEMENTS

The Second Prize Scholarship, awarded by Shanghai Jiao Tong University	2020 - 2023
Outstanding Graduates , awarded by Donghua University	Summer 2020
Academic Excellence Prize, awarded by Donghua University	Fall 2019
National Second Prize (Top 10%), the 13th Chinese University Students Intelligent Car Race	Fall 2018
Outstanding Engineer Scholarship, awarded by Donghua University	Fall 2018

SKILLSET

Programming Languages: Python, C/C++, JavaScript.

Tools & Platforms: Pytorch, Hugging Face, Docker, Matlab, Git, Vue, 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).