

Yan Tai

(+86)130-2547-4644 | taiyan2020@ia.ac.cn
github.com/MacavityT

EDUCATION

University of Chinese Academy of Sciences (GPA 3.77/4.00) Sep 2020 - Jan 2024
Artificial Intelligence Master Beijing

- **Institute of Automation, Chinese Academy of Sciences (CASIA), Supervisor: Prof. Jinqiao Wang**
- Research Interests: Multimodal LLMs, Semantic Segmentation, Object Detection, Anomaly Detection, etc.

Nanchang Hangkong University Sep 2013 - Jun 2017
Automation Engineering Bachelor Nanchang

RESEARCH EXPERIENCE

- **Yan Tai***, Weichen Fan*, Zhao Zhang, Feng Zhu, Rui Zhao, Ziwei Liu. *Link-Context Learning in Multimodal LLMs* (Submitted to AAAI2024)
 - We propose Link-Context Learning (LCL), which helps MLLMs capture novel concepts through few-shot learning.
- Bingke Zhu, **Yan Tai**, Yingying Chen, Wei Zhou, Ming Tang, Jinqiao Wang. *NextInd: Next Generation Pre-Trainer for Industrial Image Representation* (Waiting for Submission)
 - We propose NextInd, a large-scale pretrained model based on contrastive learning for industrial defect detection tasks.
- **Yan Tai**, Bingke Zhu, Yingying Chen, Ming Tang, Jinqiao Wang. *Pointrefine: Patch-attention Based Small Objects Segmentation Refiner* (Waiting for Submission)
 - We introduce a plug-and-play module that enables low-cost optimization of trained segmentation models.
- Jinqiao Wang, Yingying Chen, Bingke Zhu, **Yan Tai**, 2022. Image semantic segmentation methods, devices, electronic devices, and storage media. CN (Patent) ZL202111627261.2, filed December 29, 2021, and issued July 01, 2022.

PROFESSIONAL EXPERIENCE

SenseTime (Chinese: 商汤科技) May 2023 - Present
Research Intern Beijing

- **Link-Context Learning**: We propose LCL, emphasizing "reasoning from cause and effect" to enhance MLLM capabilities. Also, we introduce the ISEKAI dataset for evaluating recognizing new concepts. LCL outperforms other and openflamingo on ImageNet and ISEKAI. Relevant work is open-sourced and submitted to AAAI 2024. [\[Paper\]](#) [\[Code\]](#) [\[Demo\]](#)
- **MLLM Efficient Tuning**: Design MEFT, a versatile framework for MLLMs tuning. It supports various modal formats and training methods (PT/SFT/RM/DPO), along with tuning strategies like LoRA. The framework integrates conventional training features and will be open-sourced soon. [\[Code\]](#)
- Optimization of MLLMs solutions (OVD+MLLM, MLLM+ICL, etc.) for smart city tasks and reached the SOTA performance.

Institute of Automation, Chinese Academy of Sciences (CASIA) Sep 2020 - Present
Master's Student Beijing

- **RoadMainT Highway Detection**: Designed a two-stage segmentation model for precise road defect detection using self-supervision and feature fusion. The algorithm applied nationwide, achieving state-of-the-art performance.
- **Huawei Cloud Lightweight Portrait Matting**: Designed edge-preserving feathering module and lightweight segmentation model for fast portrait matting. Applied in various tasks such as meme face swapping, and so on.
- More Tasks as *Sunwoda Battery Printing Inspection*, *Oppein Furniture Board Inspection*, etc.

Aqrose Technology (Chinese: 阿丘科技) Jul 2017 - Sep 2019
Machine Learning Engineer Shenzhen

- **VIDI Development**: Implemented common image processing algorithms and modularized them into controls, supporting manual drag-and-select by users to tailor solutions for industrial detection or location tasks.
- **AIDI Development**: Implemented common AI algorithms and developed custom software mirroring VIDI's logic.

Awards

Merit Student, University of Chinese Academy of Sciences