

Yukun Zuo

Postdoctoral Research Fellow • [HomePage](#) • [LinkedIn](#) • [Google Scholar](#)
yukunzuo@umich.edu • (+1)7348829658 • WeChat: braverthanyoubelieve

Education

University of Science and Technology of China	BEIJING, CHINA
Ph.D. in Information and Communication Engineering	Sept. '18 – Jun. '23
Advisor: Prof. Changsheng Xu, A/Prof. Liansheng Zhuang	
University of Science and Technology of China	HEFEI, ANHUI, CHINA
B.S. in Information Security (School of Gifted Young)	Sept. '14 – Jun. '18

Work Experience

University of Michigan	MICHIGAN, UNITED STATES
Research Fellow, Department of Civil and Environmental Engineering	Sept. '23 – Present
Mentor: Prof. Henry X. Liu	
Institute of Automation, Chinese Academy of Sciences	BEIJING, CHINA
Research Intern, National Laboratory of Pattern Recognition	Jul. '19 – Jun. '23
Mentor: Prof. Changsheng Xu, A/Prof. Hantao Yao	

Research Interests

Representation Learning: Model Distillation, Continual Learning, Domain Adaptation. My research focuses on representation learning, particularly in continual learning and domain adaptation, with practical applications in model distillation. During my Ph.D., I concentrated on developing methods to enable models to learn more generalized and robust features, ensuring their adaptability across different data distributions and new tasks without forgetting previous knowledge. Currently, as a research fellow, I work on model distillation, transferring knowledge from large vision models to real-time models. This involves selecting important samples based on feature properties and output attributes to retrain and improve roadside perception models. I enjoy developing practical algorithms in representation learning to enhance model performance in real-world applications.

Publications

- Hierarchical Augmentation and Distillation for Class Incremental Audio-Visual Video Recognition.
Yukun Zuo, Hantao Yao, Liansheng Zhuang, Changsheng Xu.
IEEE Transactions on Pattern Analysis and Machine Intelligence (TPAMI) 2024.
- Incremental Audio-Visual Fusion for Person Recognition in Earthquake Scene.
Sisi You, **Yukun Zuo**, Hantao Yao, Changsheng Xu.
ACM Transactions on Multimedia Computing, Communications and Applications (TOMM) 2023.
- Dual Structural Knowledge Interaction for Domain Adaptation.
Yukun Zuo, Hantao Yao, Liansheng Zhuang, Changsheng Xu.
IEEE Transactions on Multimedia (TMM) 2023.
- Seek Common Ground While Reserving Differences: A Model-Agnostic Module for Noisy Domain Adaptation.
Yukun Zuo, Hantao Yao, Liansheng Zhuang, Changsheng Xu.
IEEE Transactions on Multimedia (TMM) 2022.
- Margin-based Adversarial Joint Alignment Domain Adaptation.
Yukun Zuo, Hantao Yao, Liansheng Zhuang, Changsheng Xu.
IEEE Transactions on Circuits and Systems for Video Technology (TCSVT) 2022.

Attention-Based Multi-Source Domain Adaptation.

Yukun Zuo, Hantao Yao, Changsheng Xu.

IEEE Transactions on Image Processing (TIP) 2021.

Category-Level Adversarial Self-Ensembling for Domain Adaptation.

Yukun Zuo, Hantao Yao, Changsheng Xu.

IEEE International Conference on Multimedia & Expo (ICME) 2020.

Representation Learning of Knowledge Graphs with Entity Attributes and Multimedia Descriptions.

Yukun Zuo, Quan Fang, Shengsheng Qian, Xiaorui Zhang, Changsheng Xu.

IEEE International Conference on Multimedia Big Data (BigMM) 2018.

Publication Under Review

Hierarchical Prompt for Rehearsal-free Continual Learning.

Yukun Zuo, Hantao Yao, Liansheng Zhuang, Changsheng Xu.

In Submission to IEEE Transactions on Pattern Analysis and Machine Intelligence (TPAMI) 2024.

Generalized Domain Adaptation with Extremely Limited Labels.

Yukun Zuo, Hantao Yao, Liansheng Zhuang, Changsheng Xu.

In Submission to IEEE Transactions on Circuits and Systems for Video Technology (TCSVT) 2024.

Engineering Projects

Perception System Development for Autonomous Driving

Sept. '23 – present

- **Real-Time Roadside Perception System Deployment (Dec. '23 – present).** Deploying a real-time roadside perception system integrates roadside sensors and edge devices for vehicle and pedestrian detection. The system uses V2X communication to stream data to the cloud for processing. Redis is used to manage source nodes (sensor data collection), processing nodes (object detection, tracking, and prediction), and sink nodes (data upload to cloud or local storage). Redis's pub/sub mechanism enables dynamic implementation of various functions to meet different needs.
- **On-Board Perception System Development (Sept. '23 – Apr. '24).** Developing an on-board perception system for autonomous vehicles. The system integrates multiple sensors, such as LiDAR, cameras, and GPS, along with advanced algorithms to collect, process, and interpret real-time environmental data. It includes data collection using the Lincoln MKZ open CAV platform, sensor calibration, object detection, localization, and tracking.

Research Projects

Practical Representation Learning: Model Distillation

Apr. '24 – present

- **Real-Time Roadside Perception System Improvement.** Enhancing real-time roadside perception models (YOLOv8) using unlabelled images collected over a year from various intersections in Ann Arbor. Using a large vision model (Co-DERT) for initial labelling, the project focuses on selecting important samples based on feature properties and output attributes to improve real-time roadside perception model performance, specifically focusing on Vulnerable Road Users (VRUs) detection.

Robust Representation Learning: Continual Learning

Nov. '21 – Sept. '23

- **Memorize and replay past task knowledge with prompt.** We present a novel rehearsal-free Hierarchical Prompts (H-Prompts) paradigm to address the catastrophic forgetting of prompt. H-Prompts consists of class prompt, task prompt, and general prompt to preserve the knowledge in each class, capture the past task knowledge as well as current task knowledge, and learn generalized knowledge, respectively. (In submission)
- **Exploit hierarchical structure in model and video for knowledge preservation.** We focus on class incremental audio-visual recognition, and propose a Hierarchical Augmentation and Distillation (HAD) framework considering the hierarchical structure in model (low-level and high-level) and video (snippet-level, video-level, and domain-level). HAD utilizes hierarchical data augmentation and hierarchical model distillation to conduct data knowledge preservation and model knowledge preservation, respectively. (Accepted by TPAMI 2024)

- **Ideal Annotation**

Align joint distribution and enlarge feature discrimination. We propose a Margin-based Adversarial Joint Alignment (MAJA) framework consisting joint alignment module and margin-based generative module. The joint alignment module aligns the source and target domains using the joint distribution of features and categories. The margin-based generation module generates virtual samples to conduct margin-based adversarial learning with real samples for increasing the discrimination of category features. (Accepted by TCSVT 2022)

Obtain robust pseudo-labels with structural knowledge. We design a Dual Structure Knowledge Interaction (DSKI) framework to solve the problem that the pseudo-labels generated by source-guided classifier are source-biased. We consider cluster-based knowledge and locality-based knowledge to formulate a novel structural knowledge, which assigns target-oriented and manifold-guided pseudo-labels for unlabeled target samples. In addition, we utilize two dual networks for knowledge interaction and instance mutual information maximization to learn discriminative and aligned features. (Accepted by TMM 2023)

- **Limited Annotation**

Design cross-domain clustering algorithm. We illustrate a novel Cross-Domain Knowledge Interaction (CDKI) framework to overcome the overfitting problem caused by limited annotations. Specifically, CDKI designs a new Cross-Domain Clustering (CDC) algorithm to select the confident and transferable unlabeled samples to assign pseudo labels for self-training. Moreover, a Cross-Domain Semantic Contrastive Learning (CDSCL) module is used for domain alignment by considering the category-wise semantic structure. (In submission)

- **Noisy Annotation**

Filter noisy samples with symmetrical models interaction. We introduce a Seek Common Ground While Reserving Differences (SCGWRD) strategy to reduce the impact of noisy samples utilizing the outputs of two symmetrical domain adaptation models. Specifically, Seek Common Ground strategy assumes that confident samples with consistent predictions are more likely to be clean samples, and utilizes them to implement self-training in each model. Furthermore, Reserving Differences strategy picks the confident samples with inconsistent predictions to avoid two models being converged via mutual learning. (Accepted by TMM 2022)

- **Multi-source Annotation**

Consider domain correlation to alleviate negative transfer. We build an Attention-Based Multi-Source Domain Adaptation (ABMSDA) framework by considering the domain correlations to alleviate the effects caused by dissimilar domains. Specifically, ABMSDA first trains a domain recognition model to measure domain correlations between source and target domains, and then applies the obtained domain correlations to the loss of feature alignment and supervised learning. (Accepted by TIP 2021)

Multimodal Representation Learning: Knowledge Graphs

Dec. '17 – Jun. '18

- **Embody Entity Attributes and Multimedia Descriptions for knowledge graph.** We study the problem of knowledge graph. In contrast to previous methods that only consider relationships between entities, we additionally consider entity attributes and multi-modal descriptions to learn knowledge representations, resulting in more effective performance. (Accepted by BigMM 2018)

Award

Outstanding doctoral graduates, USTC	2023
Shenzhen Stock Exchange Scholarship, USTC	2022
National Encouragement Scholarship, USTC	2016
Excellent Student Scholarship, USTC	2014, 2016

Skills

Software Programming: Python, Matlab, C/C++
Deep Learning Platform: PyTorch, Tensorflow