YIFAN WANG

Boston, MA | wang.yifan25@northeastern.edu | Google scholar

EDUCATION

Northeastern University, Boston, MA *Sep. 2023 – Now*
Ph.D in Computer Enginneering under supervision of Prof. Raymond Fu
Research Interests: Model Efficiency, Large Multi-modal model, Computer vision

Shandong University, Jinan, Shandong, P.R.China *Sep. 2018 – Jun. 2022*
Bachelor in Mathematics and applied mathematics
Core GPA: 86.3/100
Awards: 2020 Third Academic Scholarship
Core courses: Mathematical Analysis, Ordinary Differential Equations, Partial Differential Equations, Complex Variables Functions, Probability Theory, Advanced Algebra, Signals and System Analysis.

RESEARCH INTERESTS

Video Generation, Model Efficiency and Large Multimodal Model

PUBLICATIONS

Conferences:

- **Yifan Wang**, Xu Ma, Yitian Zhang, Zhongruo Wang, Sung Cheol Kim, Vahid Mirjalili, Visya Renganathan, Yun Fu. "GmNet: Revisiting Gating Mechanisms from A Frequency View". **Accepted** by ***ICLR,*** *2026*
- **Yifan Wang**, Yanyu Li, Sergey Tulyakov, Yun Fu, Anil Kag. "Diffusion-DRF: Differentiable Reward Flow for Video Diffusion Fine-Tuning". *Under review*
- **Yifan Wang**, Yun Fu. "Understanding, Abstracting and Checking: Evoking Complicated Multimodal Reasoning in LMMs". *Under review*
- **Yifan Wang**, Yi Xu, Yue Bai, Yun Fu. "Hyperbolic Few-shot Open-set Recognition". Under review

Journals:

- **Yifan Wang**, Lin Zhang, Ran Song, Lin Ma, Wei Zhang. "Exploiting Inter-Sample Affinity for Knowability-Aware Universal Domain Adaptation". **Accepted** by *International Journal of Computer Vision (**IJCV**), 2023.* [Link]
- Lin Zhang, Ran Song, **Yifan Wang**, Qian Zhang, Xiaolin Wei, Lin Ma, Wei Zhang. "Dual-Graph Contrastive Learning for Unsupervised Person Re-Identification". **Accepted** by *IEEE Transactions on Multimedia (**TMM**), 2023* [Link]
- Lin Zhang, **Yifan Wang**, Ran Song, Mingxin Zhang, Xiaolei Li, Wei Zhang. "Neighborhood-Aware Mutual Information Maximization for Source-Free Domain Adaptation" **Accepted** by *IEEE Transactions on Multimedia (**TMM**), 2023* [Link]

RESEARCH EXPERIENCE

Smile Lab@Northeastern University, *Ph.D student*, Boston, MA Sep. 2023 - Now
Advisor: Prof. Yun Raymond Fu

- GmNet: Revisiting Gating Mechanisms from A Frequency View.
  Research Brief: We systematically explore the effect of gating mechanisms on the training dynamics of neural networks from a frequency perspective. Leveraging insights on GLUs, we propose a Gating Mechanism Network (GmNet), a lightweight model designed to efficiently utilize the information of various frequency components. GmNet achieves impressive performance in terms of both effectiveness and efficiency in the image classification task.
- Hyperbolic Few-shot Open-set Recognition.
  Research brief: Propose a hyperbolic framework based on a Hyperbolic Conditional Variational Autoencoder to address the issues that existing approaches suffer from bias due to data scarcity and leading to poorly represented hierarchical class features.
- Evoking Complicated Multimodal Reasoning in LMMs.

*Research brief:*We present a prompting method named UnAC (Understanding, Abstracting, and Checking), to synergize reasoning for complicated problems in the multimodal context of LMMs, such as GPT-4V, Gemini-pro-vision and LLavA. UnAC can improve the understanding of the image capturing more details and making better reasoning of complicated multimodal questions.

Creative Vision@Snap Research, *Research Intern*, Santa Monica, CA Jun. 2025 - Now
Mentor: Anil Kag, Yanyu Li and Sergey Tulyakov

- Diffusion-DRF: Differentiable Reward Flow for Video Diffusion Fine-Tuning
  Research Brief: We propose Diffusion-DRF, a differentiable reward flow for fine-tuning video diffusion models using a frozen, off-the-shelf Vision-Language Model (VLM) as a training-free critic. Diffusion-DRF directly backpropagates VLM feedback through the diffusion denoising chain, converting logit-level responses into token-aware gradients for optimization. We propose an automated, aspect-structured prompting pipeline to obtain reliable multi-dimensional VLM feedback, while gradient checkpointing enables efficient updates through the final denoising steps.

Vsislab@Shandong University, *Research Assistant*, Jinan, Shandong, P.R.China Mar, 2021 – Mar, 2023
Advisor: Prof. Wei Zhang

- Exploiting Inter-Sample Affinity for Knowability-Aware Universal Domain Adaptation.
  Research Brief: Proposed a novel UniDA framework to fully exploring the inter-sample affinity between known and unknown samples. Specifically, we introduced a knowability-based labeling scheme which can be divided into two steps: (1) Knowability-guided detection of known and unknown samples based on the intrinsic structure of the neighborhoods of samples. (2) Label refinement based on neighborhood consistency to relabel the target samples.
- Neighborhood-Aware Mutual Information Maximization for Source-Free Domain Adaptation
  Research brief: Proposed the neighborhood-aware mutual information maximization (NAMI), which maximizes the mutual information (MI) between the representations of target domain and the corresponding neighborhood. Incorporated the consistency constraint between different views into the method to encourage the model to be robust to changes in the sample space and thus further alleviate the problem of domain gap.
- Dual-Graph Contrastive Learning for Unsupervised Person Re-Identification
  Research brief: Proposed an effective contrastive learning framework called dual-graph contrastive learning for the fully unsupervised person re-ID, which can fully utilize the intrinsic structure of the unlabeled data and better reduce the high intra-class variance. Proposed a camera-aware augmentation scheme that exploits the semantic information between samples and their corresponding camera-level centroids to relieve the domain gap between cameras.

WORK EXPERIENCE

Northeastern University, Boston, MA, USA Jun, 2024 – Aug, 2024

- **Supervisor:** Prof. Yun Raymond Fu
- **Position:** Research Assistant (Full-time)
- **Responsibilities:** Engaging the research projects in computer vision, the primary research domain is Efficient Vision Transformer.

Shandong University, Jinan, Shandong, P.R.China Mar, 2021 – Mar, 2023

- **Supervisor:** Prof. Wei Zhang
- **Position:** Research Assistant (Full-time)
- **Responsibilities:** Engaging the research projects in computer vision, the primary research domain including Domain Adaptation, Self-supervised learning and Unsupervised Learning

SKILLS AND INTERESTS

- *Computer Skills:* Python, Matlab

- *Tools and library:* Pytorch, Numpy, Scikit-learn.