Haozhe (Hank) Si

E-mail: haozhes3@illinois.edu Mobile: (+1)2172001211

EDUCATION

UNIVERSITY OF ILLINOIS AT URBANA-CHAMPAIGN

Ph.D. in Electrical and Computer Engineering   FALL 2022-SPRING 2027

GPA: 4.0

Bachelor of Science in Computer Engineering   FALL 2018-FAll 2021

GPA: 3.97

HONORS: James Scholar; Dean's List of College of Engineering (All Semester); the O. Thomas and Martha S. Purl Scholarship

RESEARCH INTERESTS: Computer Vison and Machine Learning, especially trustworthy machine learning and foundation models.

PUBLICATIONS

[1] **Provable Domain Generalization via Invariant-Feature Subspace Recovery.**

     H. Wang, **H. Si**, B. Li, H. Zhao.

     In *Proceedings of the International Conference on Machine Learning (ICML)*, 2022.

[2] **Fully Self-Supervised Depth Estimation from Defocus Clue**

     H. Si, B. Zhao, D. Wang, M. Chen, Y. Gao, Z. Wang, X. Li.

     In *Proceedings of the Computer Vision and Pattern Recognition (CVPR)*, 2023.

[3] **Enhancing Compositional Generalization via Compositional Feature Alignment**

     H. Wang*, **H. Si***, H. Shao, H. Zhao.

     Under Review of ICLR 2024

EXPERIENCE

SHANGHAI ARTIFICIAL INTELLIGENCE LAB   MAY 2023-SEPETEMBER 2023

- Researched on fast adaption of pretrained Vision-Language Models to unseen domains on visual grounding tasks; Proposed a finetuning paradigm that transfers open-world knowledge and performs UVA-VG tasks.
- Established the first benchmark for visual grounding tasks on UAV dataset; Designed a data engine for enriching the diversity of the dataset in a human-in-the-loop style. Provide a solid baseline for the new dataset.

SHANGHAI ARTIFICIAL INTELLIGENCE LAB   MARCH 2022-AUGUST 2022

- Researched on depth estimation from defocus clue; Proposed a self-learning framework that estimate depth from sparse focal stack; Design the hardware that can capture the focal stack; Address the problems exist in previous works.
- Tested the framework on NYUv2 dataset and MobileDFD dataset; Proved the method to be SOTA on depth estimation from defocus clues tasks; Verified the ability of the model to be applied in industry.
- Researched on surface reconstruction from defocus clue; Proposed a framework that incorporates defocus rendering, NeRF and surface estimation to achieve self-supervised surface reconstruction; Designed the benchmark dataset for this task.

VOKATECH   MAY 2021-OCTOBER 2021

- Researched on 3D reconstruction from few-shot image; Reproduced SOTA works including pixelNeRF, DeepSDF and PIFu and performed ablation studies on them; Analyzed the effectiveness of position encoding, transformer, and diet loss in improving reconstruction quality.
- Proposed a network architect with a novel global feature extractor that can make use of the semantic information of input images to improve the consistency of output 3D mesh; Tested on the NMR and the DTU datasets.

RESEARCH PROJECTS

LEARNING OPTIMAL REPRESENTATIONS FOR COMPOSITIONAL GENERALIZATION   SPRING 2023

- Designed a *Finetuning with Multi-label Classification* paradigm that encourage the feature extractors learning the representations

that is domain invariant. Adopted the proposed paradigm to CLIP and DINOv2. Boosted their performance on various o.o.d. dataset.

- Developing a ViT-Finetune package that aims helping users finetune various ViT models easily and efficiently, e.g., finetune a CLIP ViT-G/14 model (1.8B parameters, needs 512 A100s in pre-training) on a single GPU.

BYZANTINE-ROBUST FEDERATED DOMAIN GENERALIZATION   FALL 2022

- Designed a scenario in federated learning where both OOD data and poisoned data exist in the clients; Showed the SOTA byzantine-robust federated learning algorithm will fail to generalize in such scenario by being too conservative.
- Proposed a novel federated learning algorithm that encourages the features to align across the domains; the proposed method makes trade-off between domain generalization and byzantine robustness.

ISR: INVARIANT SUBSPACE RECOVERING   SUMMER 2021

- Worked with Professor Han Zhao and Professor Bo Li on domain generalization; Purposed a novel post-processing method to recover the invariant subspace of features and improve the robustness of models; The work is published on ICML 2022.
- Tested the algorithm on both synthetic and real-world image datasets for domain generalization; Outperformed the algorithms including IRM and Group-DRO.

PACB: PROBLEM AGNOSTIC CLUSTER-BASED AUDIO PRETRAINING   FALL 2021

- Proposed a novel self-supervised training scheme to better leverage the large corpus of unlabeled audio data; Designed a problem-agnostic cluster-based pretext task to pretrain the feature extractor.
- Tested the scheme on source separation and speaker classification datasets; the models converge faster and result in higher accuracy in both tasks.

FEDERATED DOMAIN GENERALIZATION   FALL 2021

- Proposed a novel federated learning scheme making use of the data distribution of public and private data and can handle the out of distribution data on the clients.
- Designed the training framework for the proposed scheme; Designed a problem setting on CelebA and constructed the dataset for federated learning that requires domain generalization; Outperformed the FedAvg optimizer on both ID and OOD accuracy.

HAODE-NET: OBJECT DETECTION IN HAZY ENVIRONMENT   SPRING 2021

- Utilized continual learning, multitask learning and knowledge distillation to propose a novel network architecture targeting object detection in a hazy environment; Tested on the SSD and the RTTS datasets.
- Performed extensive experiments and ablation studies verifying the superiority of the proposed framework.

ACDEMICRANK   SPRING 2021

- Worked in the FORWARD Data Lab with Professor Kevin Chang; Conducted research on calculating the rank of academic works about a given keyword according to their importance.
- Formalized the task into a weighted Personalized PageRank problem with multiple nodes values; Re-processed 3 datasets for the task to shrink the data size and speed up calculation.
- Designed a fast-ranking algorithm that works on citation relation graphs that can calculate the ranking of a graph with more than 10 billion edges in about 1 hour; Generated convincing results.

SKILLS

DEEP LEARNING FRAMEWORKS

- PyTorch; Scikit-learn; Jax; TensorFlow;

PROGRAMMING LANGUAGE

- Python; C/C++; LaTex; SQL; System Verilog; x86 Assembly; Clojure; HTML; MATLAB;

APPLICATIONS

- Microsoft Office; Photoshop; Premiere Pro; Autodesk; V-REP; ROS.