Zhiheng Li

: https://zhiheng.li : zhiheng.li@rochester.edu : +1-585-351-1179
: https://www.linkedin.com/in/zhiheng-li-ur/ : https://github.com/zhihengli-UR

Education

University of Rochester Rochester, NY
Ph.D. in Computer Science Aug 2018 - Present
Advisor: Chenliang Xu

Wuhan University Wuhan, China
B.Eng. in Software Engineering Sep 2014 - June 2018

Research Interests

`General Research Interests`: Computer Vision and Trustworthy AI (i.e., Responsible AI)
`Specific Research Interests`: AI Fairness; Out-of-Distribution Robustness; Bias Detection and Bias Mitigation

Internship Experience

Meta AI Seattle, WA
Research Intern May 2022 - Dec 2022
Project: **Multi-shortcut Mitigation** Mentors: Ivan Evtimov,Mark Ibrahim, Albert Gordo

- `Problem Setup`: While previous works have the tenuous assumption that only a single shortcut (i.e., bias) exists, images are rife with multiple visual cues. Thus, we study a neglected but critical problem—multi-shortcut learning problem in image classification.
- `New Datasets`: We create two datasets—(1) *UrbanCars*, a semi-synthetic dataset with multiple precisely controlled spurious cues; (2) *ImageNet-W*, a new out-of-distribution variant of ImageNet to evaluate the reliance on *watermark*—a new shortcut that we discovered in ImageNet, which affects a broad range of vision models, including supervised ResNet, self-supervised models (e.g., MoCoV3 and MAE), and even large foundation models (e.g., SEER, SWAG, and CLIP).
- `New Finding`: Benchmarking on both datasets reveals the overlooked challenge in shortcut learning that resembles a Whac-A-Mole game—mitigating one shortcut amplifies others.
- `Method`: We propose Last Layer Ensemble (LLE) based on data augmentation to mitigate multiple shortcuts jointly.
- `Results`: Our proposed LLE method does not suffer the Whac-A-Mole problem.
- `Publication`: The paper [1] is submitted to a top-tier computer vision conference in 2023.

NEC Labs America Princeton, NJ
Research Intern Jun 2021 - Sep 2021
Project: **Compositional Text-to-Image Synthesis** Mentors: Martin Renqiang Min,Kai Li

- `Problem Setup`: Existing text-to-image methods lack compositionality. E.g., They struggle to generate images of "*a man wearing lipstick*," leading to fairness and robustness issues.
- `Method`: We propose StyleT2I, a novel text-to-image synthesis framework. To improve compositionality, we propose: (1) CLIP-guided contrastive loss; (2) leveraging disentangled representations in StyleGAN to control image generation.
- `Results`: StyleT2I outperforms previous methods, e.g., ControlGAN and TediGAN, on CelebA-HQ and CUB datasets.
- `Publication`: The paper [2] is accepted to CVPR 2022.

Research Projects at University of Rochester

Project: **Unknown Bias Discovery and Mitigation** Apr 2021 - Feb 2022

- `Problem Setup`: Biases in datasets are usually unknown in terms of (1) bias type; and (2) number of biases. Even for known bias, bias labels (e.g., gender labels) are hard to collect due to ethical and privacy concerns. The goal is to discover and mitigate unknown biases without any prior knowledge of bias types or numbers of biases.
- `Method`: We propose DebiAN, which jointly trains two networks, i.e., a classifier and a discoverer, in an alternate fashion to discover and mitigate unknown biases.
- `New Dataset`: We construct Multi-Color MNIST, a new dataset containing multiple biases, to study the mitigation of multiple biases, which is underexplored in works.
- `Results`: DebiAN outperforms existing approaches, e.g., LfF and EIIL, on CelebA, BAR, bFFHQ, and Places datasets. Besides, our Multi-Color MNIST dataset demonstrates DebiAN's advantage of mitigating multiple biases, whereas LfF, a previous seminal work, can only identify and mitigate one bias. Besides, DebiAN discovers interesting unknown biases, e.g., visible hair area for gender classification and indoor vs. outdoor scene for restaurant class in scene classification.
- `Publication`: The paper [3] is accepted to ECCV 2022.

Project: **Unknown Bias Discovery** Sep 2020 - Mar 2021

- `Problem Setup`: Existing bias detection frameworks heavily rely on human assumptions of the bias type. As a result, unknown bias cannot be discovered, not to mention to be mitigated. Thus, we formalize a new task, Unknown Bias Discovery.
- `Method`: Leveraging the disentangled latent space of generative models (e.g., StyleGAN), we optimize the latent hyperplane that represents an unknown bias.
- `Results`: Our method discovers some interesting unknown biases in ImageNet dataset (e.g., the shade of fur color for cat) and Places dataset (e.g., Eiffel tower for towel class in scene classification).
- `Publication`: The paper [4] is accepted to ICCV 2021.

Project: **Interpretable Visual Representation via Perceptual Grouping on CNN** Sep 2020 - Mar 2021

- `Problem Setup`: The commonly used convolutional neural nets (CNN) fall short in interpretability and efficiency. We address the limitations by incorporating the traditional perceptual grouping process into CNNs.
- `Method`: We propose Deep Group Model (DGM), which performs grouping hierarchically over CNN feature maps.
- `Results`: DGM outperforms existing approaches in unified perceptual parsing task, which consists of image-level scene classification and multiple pixel-level segmentation tasks, e.g., object segmentation, part segmentation, etc. Besides, DGM reduces computational overhead compared to pixel-based self-attention approaches. Moreover, DGM facilitates explainable CNN by generating saliency maps with sharp boundaries that locate spatial regions more precisely.
- `Publication`: The paper [5] is accepted to CVPR 2020.

Project: **Lip Movements Generation** Sep 2020 - Mar 2021

- `Problem Setup`: Given audio and an input lip image, the task is to generate a lip movements video of the mouth speaking the voice in the given audio.
- `Method`: Based on conditional GAN, we propose novel losses to model the correlation between lip movements and audio.
- `Results`: The generated videos have good visual quality. The generated lip movements and audio are in good synchronization, facilitating many future works on the talking head generation task.
- `Publication`: The paper [6] is accepted to ECCV 2018.

Publication

[1] **Zhiheng Li**, Ivan Evtimov, Albert Gordo, Caner Hazirbas, Tal Hassner, Cristian Canton Ferrer, Chenliang Xu, and Mark Ibrahim. A Whac-A-Mole Dilemma: Shortcuts Come in Multiples Where Mitigating One Amplifies Others. In *In Submission*, 2023.

[2] **Zhiheng Li**, Martin Renqiang Min, Kai Li, and Chenliang Xu. StyleT2I: Toward Compositional and High-Fidelity Text-to-Image Synthesis. In *CVPR*, 2022.

[3] **Zhiheng Li**, Anthony Hoogs, and Chenliang Xu. Discover and Mitigate Unknown Biases with Debiasing Alternate Networks. In *ECCV*, 2022.

[4] **Zhiheng Li** and Chenliang Xu. Discover the Unknown Biased Attribute of an Image Classifier. In *ICCV*, 2021.

[5] **Zhiheng Li**, Wenxuan Bao, Jiayang Zheng, and Chenliang Xu. Deep Grouping Model for Unified Perceptual Parsing. In *CVPR*, 2020.

[6] Lele Chen*, **Zhiheng Li* (*Equal Contribution)**, Ross K Maddox, Zhiyao Duan, and Chenliang Xu. Lip Movements Generation at a Glance. In *ECCV*, 2018.

[7] Raghav Mehta, Vítor Albiero, Li Chen, Ivan Evtimov, Tamar Glaser, **Zhiheng Li**, and Tal Hassner. You Only Need a Good Embeddings Extractor to Fix Spurious Correlations. In *ECCV Responsible Computer Vision Workshop*, 2022.

[8] Jie Chen, **Zhiheng Li**, Jiebo Luo, and Chenliang Xu. Learning a Weakly-Supervised Video Actor-Action Segmentation Model with a Wise Selection. In *CVPR*, 2020.

[9] **Zhiheng Li**, Geemi P. Wellawatte, Maghesree Chakraborty, Heta A. Gandhi, Chenliang Xu, and Andrew D. White. Graph Neural Network Based Coarse-Grained Mapping Prediction. *Chemical Science*, 2020.

Professional Services

- **Reviewer**: NeurIPS'20, CVPR'21, ICML'21, ICCV'21, NeurIPS'21, ICLR'22, AAAI'22, CVPR'22, ICML'22, ECCV'22, NeurIPS'22 (top reviewer), CVPR'23, TMLR
- **Volunteer**: FAccT'21, ICLR'21

Skills

- **Programming Languages**: Python, Java, C, C++, Swift, Objective-C, Ruby
- **Frameworks**: PyTorch, NumPy, OpenCV, Caffe, MxNet