Hossein Souri

hsouri.github.io hsouri1@jhu.edu github.com/hsouri LinkedIn

EDUCATION

Johns Hopkins University — **Baltimore, MD**
PhD in Computer Science — *Expected Graduation: Aug 2024*

Johns Hopkins University — **Baltimore, MD**
MS in Computer Science — *Aug 2020 - Aug 2022*

University of Maryland, College Park — **College Park, MD**
MS in Electrical and Computer Engineering — *Aug 2018 - Aug 2020*

RESEARCH INTERESTS

- Responsible and Safe AI
- Generative AI
- Multimodal ML, VLM
- AI Security, Fairness
- Adversarial Robustness
- Transfer Learning

PUBLICATIONS

15. *Generating Potent Poisons and Backdoors from Scratch with Guided Diffusion*
Hossein Souri, Arpit Bansal, Hamid Kazemi, Liam Fowl, Aniruddha Saha, Jonas Geiping, Andrew Gordon Wilson, Rama Chellappa, Tom Goldstein, Micah Goldblum
Oral presentation. ICML 2024 Workshop on the Next Generation of AI Safety [Paper][Code]

14. *Identifying Attack-Specific Signatures in Adversarial Examples*
Hossein Souri, Pirazh Khorramshahi, Chun Pong Lau, Micah Goldblum, and Rama Chellappa
IEEE ICASSP 2024 [Paper][Code]

13. *Battle of the Backbones: A Large-Scale Comparison of Pretrained Models across Computer Vision Tasks*
Micah Goldblum*, **Hossein Souri***, Renkun Ni, Manli Shu, Viraj Prabhu, Gowthami Somepalli, Prithvijit Chattopadhyay, Mark Ibrahim, Adrien Bardes, Judy Hoffman, Rama Chellappa, Andrew Gordon Wilson, Tom Goldstein
NeurIPS 2023 [Paper][Code]
In collaboration with Facebook AI Research (FAIR)

12. *A Deep Dive into Dataset Imbalance and Bias in Face Identification*
Valeriia Cherepanova, Steven Reich, Samuel Dooley, **Hossein Souri**, Micah Goldblum, Tom Goldstein
AAAI/ACM Conference on AI, Ethics, and Society (AIES) 2023 [Paper]

11. *Interpolated Joint Space Adversarial Training for Robust and Generalizable Defenses*
Chun Pong Lau, Jiang Liu, **Hossein Souri**, Wei-An Lin, Soheil Feizi, Rama Chellappa
TPAMI 2023 [Paper]

10. *Sleeper Agent: Scalable Hidden Trigger Backdoors for Neural Networks Trained from Scratch*
Hossein Souri, Liam Fowl, Rama Chellappa, Micah Goldblum, and Tom Goldstein
NeurIPS 2022 [Paper][Code]

9. *Pre-Train Your Loss: Easy Bayesian Transfer Learning with Informative Priors*
Ravid Shwartz-Ziv, Micah Goldblum, **Hossein Souri**, Sanyam Kapoor, Chen Zhu, Yann LeCun, Andrew Gordon Wilson
NeurIPS 2022 [Paper][Code]

8. *Thinking Two Moves Ahead: Anticipating Other Users Improves Backdoor Attacks in Federated Learning*
Yuxin Wen, Jonas Geiping, Liam Fowl, **Hossein Souri**, Rama Chellappa, Micah Goldblum, Tom Goldstein
ICML 2022 Workshop on AdvML Frontiers [Paper][Code]

7. *Mutual Adversarial Training: Learning Together is Better Than Going Alone*
Jiang Liu, Chun Pong Lau, **Hossein Souri**, Soheil Feizi, Rama Chellappa
IEEE Transactions on Information Forensics and Security (TIFS) 2022 [Paper]

6. *The Close Relationship Between Contrastive Learning and Meta-Learning*
Renkun Ni, Manli Shu, **Hossein Souri**, Micah Goldblum, Tom Goldstein
ICLR 2021 [Paper][Code]

5. *ATFaceGAN: Single Face Image Restoration and Recognition from Atmospheric Turbulence*
Chun Pong Lau, **Hossein Souri**, Rama Chellappa
IEEE International Conference on Automatic Face and Gesture Recognition (FG) 2020
Best paper (honorable mention) award. **Oral presentation** [Paper]

Book Chapters

4. *Adversarial Attacks and Robust Defenses in Deep Learning*
Chun Pong Lau, Jiang Liu, Wei-An Lin, **Hossein Souri**, Pirazh Khorramshahi, Rama Chellappa
Elsevier 2023 [Paper]

Preprints

3. *GANs with Variational Entropy Regularizers: Applications in Mitigating the Mode-Collapse Issue*
Pirazh Khorramshahi*, **Hossein Souri***, Rama Chellappa, Soheil Feizi. [Paper]

2. *Towards Gender-Neutral Face Descriptors for Mitigating Bias in Face Recognition*
Prithviraj Dhar, Joshua Gleason, **Hossein Souri**, Carlos D. Castillo, Rama Chellappa. [Paper]

1. *An Adversarial Learning Algorithm for Mitigating Gender Bias in Face Recognition*
Prithviraj Dhar, Joshua Gleason, **Hossein Souri**, Carlos D. Castillo, Rama Chellappa. [Paper]

EMPLOYMENT

Ping An Technology, Silicon Valley Research Lab **Palo Alto, CA**
ML/CV Research Scientist Intern *May 2023 - Nov 2023*

- Throughout my internship, I actively contributed to a prominent virtual being project. Specifically, I focused on the Talking Head Video Generation task, where I successfully designed and implemented a range of multi-modal generative models. By utilizing state-of-the-art diffusion models, I integrated videos, audio, and text resulting in highly realistic talking head animations.

Artificial Intelligence for Engineering and Medicine Lab (AIEM) **Baltimore, MD**
ML/CV Research Assistant *Aug 2020 - Present*

- My primary research is applied machine learning and computer vision, focusing on improving the robustness, transferability, and performance of image/face/video classifiers, object detection/segmentation, and generative models. Those include adversarial robustness, transfer learning, self-supervised learning, diffusion models, GANs, as well as, data poisoning and backdoor attacks.

University of Maryland Institute for Advanced Computer Studies **College Park, MD**
ML/CV Research Assistant *Aug 2018 - Aug 2020*

- My research was focused on Generative Adversarial Networks (GANs), image restoration, face recognition, and fairness.

TEACHING EXPERIENCE

- Teaching Assistant, Machine Intelligence, Spring 2022
- Teaching Assistant, Machine Perception, Fall 2021
- Teaching Assistant, Deep Learning, Spring 2021
- Teaching Assistant, Machine Learning, Spring 2020

Community Involvement

- Conference Reviewer
  - CVPR 2022,2023,2024
  - NeurIPS 2022,2023,2024
  - ICLR 2023,2024
  - ICML 2023,2024
  - ECCV 2022,2024
  - WACV 2022,2023,2024