Sahil Singla

Phone : +1.475.228.4315

Apt 1, 570 Fall River Ter
Sunnyvale

Email : ssingla@umd.edu
Web: singlasahil14.github.io/

Research interests

Explainability and Reliability aspects of deep learning

Employment

- **Google Research** Mountain View, California
  Research Scientist *January 2023 - Present*

Publications

- **Sahil Singla**, Soheil Feizi. **Salient Imagenet, How to discover spurious features in deep learning?**. Accepted at **ICLR, 2022**.
  https://arxiv.org/abs/2110.04301

- **Sahil Singla**, Surbhi Singla, Soheil Feizi. **Improved deterministic l2 robustness on CIFAR-10 and CIFAR-100**. Accepted at **ICLR, 2022 (Spotlight, top 4% submissions)**.
  https://openreview.net/forum?id=tD7eCtaSkR

- **Sahil Singla**, Besmira Nushi, Shital Shah, Ece Kamar, Eric Horvitz. **Understanding Failures of Deep Networks via Robust Feature Extraction**. Accepted at **CVPR, 2021 (Oral, top 4% submissions)**.
  https://arxiv.org/abs/2012.01750

- Cassidy Laidlaw, **Sahil Singla**, Soheil Feizi. **Perceptual Adversarial Robustness: Defense Against Unseen Threat Models** Accepted at **ICLR, 2021**.
  https://openreview.net/forum?id=dFwBosAcJkN

- **Sahil Singla**, Soheil Feizi. **Fantastic Four: Differentiable and Efficient Bounds on Singular Values of Convolution Layers.** Accepted at **ICLR, 2021**.
  https://openreview.net/forum?id=JCRblSgs34Z

- **Sahil Singla**, Soheil Feizi. **Skew Orthogonal Convolutions**. Accepted at **ICML, 2021.**.
  https://arxiv.org/abs/2105.11417

- Vasu Singla, **Sahil Singla**, Soheil Feizi, David Jacobs. **Low Curvature Activations Reduce Overfitting in Adversarial Training**. Accepted at **ICCV, 2021**.
  https://arxiv.org/abs/2102.07861

- Vedant Nanda, Samuel Dooley, **Sahil Singla**, Soheil Feizi, John Dickerson. **Fairness Through Robustness: Investigating Robustness Disparity in Deep Learning.** Accepted at **FAccT (formerly FAT), 2021.**
  https://arxiv.org/abs/2006.12621

- **Sahil Singla**, Soheil Feizi. **Second-Order Provable Defenses against Adversarial Attacks**. Accepted at **ICML, 2020**.
  https://arxiv.org/abs/2006.00731

- **Sahil Singla**, Eric Wallace, Shi Feng, Soheil Feizi. **Understanding Impacts of High-Order Loss Approximations and Group Features in Interpretation.** Accepted at **ICML, 2019**.
  https://arxiv.org/abs/1902.00407

- **Sahil Singla**, Soheil Feizi. **Improved techniques for deterministic l2 robustness**. Accepted at **NeurIPS, 2022**.

- Mazda Moayeri, **Sahil Singla**, Soheil Feizi. **Hard ImageNet: Segmentations for Objects with Strong Spurious Cues**. Accepted at **NeurIPS, 2022**.
- Mazda Moayeri, Wenxiao Wang, **Sahil Singla**, Soheil Feizi. **Spuriosity Rankings: Sorting Data to Measure and Mitigate Biases** . Accepted at **NeurIPS, 2023**.
- **Sahil Singla**, Atoosa Chegini, Mazda Moayeri, Soheil Feizi. **Data-Centric Debugging: mitigating model failures via targeted image retrieval** . Accepted at **WACV, 2024**.

Education

- **University of Maryland** College Park, MD
  PhD. Research advisor: Prof. Soheil Feizi *Aug. 2018 – Dec. 2022*
- **Indian Institute of Technology, Delhi** New Delhi, India
  Bachelor of Technology in Computer Science *Aug. 2010 – July. 2014*

Research Internships

- **Microsoft Research** Redmond, Washington
  Worked with Besmira Nushi, Ece Kamar, Shital Shah, Eric Horvitz *June 2020 - August 2020*
  - Worked on failure explanation of deep neural networks using robustness
  - Paper accepted in CVPR 2021 titled ”Understanding Failures of Deep Networks via Robust Feature Extraction”

Invited Talks

- **London Machine Learning Meetup** Online
  Salient Imagenet: How to discover spurious features in deep learning? *16 February 2022*
- **Stanford, AI for Medical Imaging (AIMI) center** Stanford, California
  Understanding Failures of Deep Networks via Robust Feature Extraction *10 June 2021*
- **Microsoft Research, ASI Group** Redmond, Washington
  Visual feature extraction for error analysis *14 August 2020*
- **Microsoft Research, MLO Group** Redmond, Washington
  Second-Order Provable Defenses against Adversarial Attacks *22 July 2020*

Awards and Academic achievements

- **Outstanding Research Assistant Award**. Awarded to top 2% graduate research assistants every year by the Graduate School at the University of Maryland.
- **Dean’s Fellowship**. Cash prize of $2500. Awarded to only two students in the first and second year in the Computer Science department at University of Maryland.
- Secured **All India Rank 47** out of half a million students (amongst top .01% of the students) who appeared in **IIT-JEE 2010** exam
- Secured **All India Rank 56** out of one million students (amongst top .005% of the students) in **AIEEE-2010** exam

Prior Work Experience

- **Goldman Sachs** Bangalore, India
  Analyst *August 2014 - August 2015*
  - Worked on reducing the time taken for pricing options.
  - Developed a software to calculate various risks associated with options portfolio
- **WaltonPay** New Delhi, India
  Cofounder and CTO *August 2015 - March 2016*

  - Developed a mobile app that would gather SMS data for credit evaluation.
  - Designed a statistical model to evaluate a persons credit profile based on SMS data.

- **Farmguide** Gurgaon, India
  Machine Learning Engineer *April 2016 - March 2017*
  - Developed a software to segment farm boundaries from satellite imagery
  - Work was featured in Forbes and is currently being used by Government of India

- **APUS** Gurgaon, India
  Machine Learning Engineer *April 2017 - July 2017*
  - Implemented neural style transfer that runs faster than popular app Prisma on phone.
  - Implemented the tensorflow op for sparse convolution in C++ that can run on mobile phone.

- **Computer Vision Consulting** Gurgaon, India
  Consultant *August 2017 - December 2018*
  - Use satellite imagery to identify areas of low and high agriculture produce.
  - Use computer vision to estimate weight of agriculture produce in a container.

- **Quadeye Securities** Gurgaon, India
  Quantitative Analyst *Jan 2018 - August 2018*
  - Designed a machine learning model to predict whether to buy/sell based on analyst ratings.
  - Designed a statistical model to reduce the runtime of an algorithm for strategy optimization.

References

- Soheil Feizi
  - Assistant Professor, University of Maryland, College Park
  - Email: sfeizi@cs.umd.edu
- Eric Horvitz
  - Chief Scientific Officer, Microsoft Research
  - Email: horvitz@microsoft.com
- David Jacobs
  - Professor, University of Maryland, College Park
  - Email: djacobs@cs.umd.edu