Anirudh Sundara Rajan

asundararaj2@wisc.edu in anirudhsundar Google Scholar GitHub

EDUCATION

University of Wisconsin-Madison **2022–Present**
MS/PhD in Computer Science GPA: 3.841/4
- *Thesis: What Knowledge Gets Distilled in Knowledge Distillation?*

BITS Pilani **2018–2022**
B.E. in Electronics and Communication Engineering GPA: 8.447/10

RESEARCH EXPERIENCE

- **Research Assistant**, *University of Wisconsin-Madison* Jan 2023 – Present
  Advisor: Prof. Yong Jae Lee
  - Conducted in-depth study on Knowledge Distillation. Work accepted to **NeurIPS 2023**.
  - Worked on image editing using Stable Diffusion.
  - Currently working on image forensics with a focus on Diffusion models.
- **Research Intern**, *The Artificial Intelligence Institute of South Carolina* June 2021 – May 2022
  Advisor: Dr. Amit Sheth
  - Knowledge Infused clustering of cooking representations published in **Frontiers**.
  - Contributed to constructing a biomedical knowledge graph.

PUBLICATIONS & PREPRINTS

* indicates equal contribution.

- **arXiv 2025: Rajan, A. S.**., Lee, Y. J. (2025). *Stay-Positive: A Case for Ignoring Real Image Features in Fake Image Detection*. arXiv:2502.07778.
- **ICLR 25: Rajan, A. S***., Ojha, U*., Schloesser, J., Lee, Y. J. (2024). *Aligned Datasets Improve Detection of Latent Diffusion-Generated Images.* arXiv preprint arXiv:2410.11835.
- **NeurIPS 23:** Ojha, U*., Li, Y*., **Sundara Rajan, A***., Liang, Y., Lee, Y. J. (2023). *What knowledge gets distilled in knowledge distillation?* Advances in Neural Information Processing Systems, 36, 11037-11048.
- **Frontiers 23:** Venkataramanan, R., Padhee, S., Rao, S. R., Kaoshik, R., **Sundara Rajan, A.**, Sheth, A. (2023). *Ki-Cook: clustering multimodal cooking representations through knowledge-infused learning.* Frontiers in Big Data, 6, 1200840.

RESEARCH PROJECTS

Mitigating Spurious Correlations in Fake Image Detection **Details**
- Identified pitfalls in existing methods that rely on compression and resizing artifacts for detecting images generated by latent diffusion models.
- Leveraged the Stable Diffusion autoencoder to reconstruct real images into fake images, demonstrating that training a detector on this dataset mitigates spurious correlations.
- Proposed method outperforms the previous SOTA, achieving a +1.53 AP improvement for Stable Diffusion-generated images and +3.49 AP for Midjourney-generated images.

Improved AEROBLADE for Detection of Post-Processed Images **Details**
- Highlighted limitations of an existing reconstruction-based fake image detection algorithm, which fails when images undergo post-processing such as compression/resizing.
- Replaced the distance-based decision-making approach with a training-based method, achieving significant improvements measured by true positive rate (TPR) @ 5 false positive rate (FPR) for detecting images from Stable Diffusion and Midjourney.
- Observed that layers 10-14 in the CLIP encoder contain useful representations for detecting fake images. Improved the AEROBLADE approach by replacing the VGG-based representations with CLIP, achieving better performance with CLIP compared to VGG.

Drag-Based Editing of Images through Energy-Based Guidance **Details**

- Proposed an inference-based technique for drag-based image editing, allowing users to select and reposition points in an image while maintaining realism.
- Found that representations captured by the second block of the Stable Diffusion UNet decoder contain rich semantic information, enabling localization of regions to be modified in images.
- Designed energy functions to reposition objects while preserving overall image details.

Socratic AI for Coding Assistance **Code**

- Developed a Flask-based Socratic Code Tutor that provides dynamic suggestions to students based on their code, test cases, and instructors' solutions using GPT-4o responses. Conducted a between-participants factorial study (n=10) using ANCOVA to provide descriptive and inferential insights on our application's effectiveness in aiding students' coding.

Contextual Representations for Fine-Grained Emotion Classification **Code**

- Analyzed the performance of large language models (e.g., BERT, RoBERTa) on fine-grained emotion classification using the GoEmotions dataset.
- Trained RoBERTa for emotion classification by fine-tuning on a small number of data points selected through an entropy-based active learning criterion and labeled using GPT-3.5

TEACHING EXPERIENCE

- **CS 839 — Learning-Based Image Synthesis, UW-Madison** *Teaching Assistant, Aug 2023 – Dec 2023, Aug 2024 – Dec 2024*
  Instructed by Prof. Yong Jae Lee.
  - Designed problem sets to assess students' understanding of deep generative modeling.
  - Delivered tutorials on PyTorch and the Google Cloud Platform.
  - Structured the course's paper reading program by curating a selection of foundational and recent papers across diffusion models, GANs, and autoregressive models, ensuring coverage of both classic and cutting-edge developments in these areas.
- **COMP SCI 220 — Data Science Programming I, UW-Madison** *Teaching Assistant, Jan 2023 – May 2023*
  Instructed by Prof. Michael Doescher.
  - Prepared assignments to assess students' Python programming proficiency.
- **COMP SCI 200 — Programming I, UW-Madison** *Teaching Assistant, Jun 2022 – Dec 2022*
  Instructed by Prof. Jim Williams.
  - Prepared assignments to assess students' Java programming proficiency.

MATHEMATICS COURSEWORK

University of Wisconsin-Madison
Introduction to Stochastic Processes
Non-Linear Optimization

BITS Pilani
Mathematics I: Vector Calculus
Mathematics II: Linear Algebra
Mathematics III: Differential Equations
Discrete Mathematics

MACHINE LEARNING COURSEWORK

University of Wisconsin-Madison
Machine Learning
Advanced Natural Language Processing
Learning Based Image Synthesis

BITS Pilani
Neural Networks and Fuzzy Logic
Data Mining

SKILLS

Programming Languages: Java, C, C++, Python
Deep Learning Frameworks: PyTorch, Transformers, Diffusers