

SAVYA KHOSLA

savyak2@illinois.edu • [LinkedIn](#) • [GitHub](#) • [Google Scholar](#) • [Personal Webpage](#)

RESEARCH INTERESTS

Representation Learning: Creating efficient and scalable visual representations that can substitute traditional patch-based encoding

Long-Form Video Understanding: Enabling fine-grained retrieval and understanding from episodic memory

Multimodal Learning: Developing unified systems that understand modalities like vision, language, audio, and even bio-signals

EDUCATION

University of Illinois Urbana-Champaign

Ph.D. in Computer Science

Urbana, IL

Aug 2024 - Present

- Advisors: Prof. Derek Hoiem and Prof. Alex Schwing

University of Illinois Urbana-Champaign

M.S. in Computer Science (CGPA: 4.0 / 4.0)

Urbana, IL

Aug 2022 - May 2024

- Advisor: Prof. Derek Hoiem
- Thesis: [Dynamic Multimodal Learning](#)

Delhi Technological University

B.Tech. in Computer Engineering (CGPA: 9.40 / 10.0)

New Delhi, DL

Aug 2017 - July 2021

- Advisor: Prof. Anil Singh Parihar

RESEARCH EXPERIENCE

Meta Reality Labs

Research Scientist Intern

Burlingame, CA

May 2025 - Oct 2025

- Working on a multimodal hand foundation model that can learn to represent hand state information like pose, force, and stiffness

University of Illinois Urbana-Champaign

Research Fellow

Urbana, IL

Aug 2024 - May 2025

- Developed REN [[NeurIPS 2025](#)], a region-based image encoder that represents each semantically meaningful region with a token
- Achieved 25× reduction in token count compared to patch encoders while outperforming them on segmentation and retrieval tasks

Adobe Research

Research Scientist Intern

San Jose, CA

May 2024 - Aug 2024

- Proposed MAGNET [[ACL 2025](#)], a unified framework that can enhance any off-the-shelf LLM with representation learning and infilling capabilities, while retaining its core generative functions

University of Illinois Urbana-Champaign

Research Assistant (Guide: Prof. Derek Hoiem)

Urbana, IL

Jan 2024 - May 2024

- Developed RELOCATE [[CVPR 2025](#)], a framework for spatiotemporal localization of small and fleeting objects in long videos
- Demonstrated accurate localization in real-world conditions, achieving 49% relative improvement over baselines on Ego4D

Allen Institute for AI

Research Intern

Seattle, WA

Oct 2022 - Dec 2023

- Contributed to Unified-IO 2 [[CVPR 2024](#)], the first autoregressive multimodal model that can understand and generate images, text, audio, and actions

National University of Singapore

Research Assistant (Guide: Prof. Kenji Kawaguchi and Alex Lamb)

Remote

Apr 2022 - Aug 2022

- Designed (1) an active learning algorithm robust to heteroskedastic noise; and (2) a self-supervised fine-tuning method to leverage unlabeled data in active learning setting, improving performance by 15% and reducing annotation costs [[ECAI 2023](#)]

Mila - Quebec AI Institute

Research Intern (Guide: Prof. Yoshua Bengio)

Montreal, QC

Apr 2021 - Nov 2021

- Exposed limitations of uncertainty-based active learning by introducing three heteroskedastic data distributions [[NeurIPS 2021](#)]
- Developed adversarial training that reduces clean-data error by 48% while preserving robustness to adversarial attacks [[INNS 2022](#)]

Delhi Technological University

Undergraduate Researcher (Guide: Prof. Anil Singh Parihar)

New Delhi, DL

Apr 2021 - Nov 2021

- Leveraged image-based representation of malware binaries and ensembling to develop state-of-the-art model for malware detection and classification [[MMTA 2022](#), [CONIT 2021](#)]

SELECTED PUBLICATIONS & PREPRINTS

A full list of publications can be seen on my [Google Scholar](#) author page.

(* denotes equal contribution, † denotes alphabetical order)

1. **REN: Fast and Efficient Region Encodings from Patch-Based Image Encoders** [Link](#)
[Savya Khosla](#), Sethuraman TV, Barnett Lee, Alex Schwing, Derek Hoiem
Neural Information Processing Systems, 2025
2. **RELOCATE: A Simple Training-Free Baseline for Visual Query Localization Using Region-Based Representations** [Link](#)
[Savya Khosla](#), Sethuraman TV, Alex Schwing, Derek Hoiem
Computer Vision and Pattern Recognition, 2025
3. **MAGNET: Augmenting Generative Decoders with Representation Learning and Infilling Capabilities** [Link](#)
[Savya Khosla](#), Aditi Tiwari, Kushal Kaffle, Simon Jenni, Handong Zhao, John Collomosse, Jing Shi
Association for Computational Linguistics, 2025
4. **FRAME: Feature Representation and Anticipation with Memory** [Link](#)
Sethuraman TV, [Savya Khosla](#), Vignesh Srinivasakumar, Jiahui Huang, Seoung Wug Oh, Simon Jenni, Derek Hoiem, Joon-Young Lee
arXiv:2506.05543, 2025
5. **Unified-IO 2: Scaling Autoregressive Multimodal Models with Vision, Language, Audio, and Action** [Link](#)
Jiasen Lu*, Christopher Clark*, Sangho Lee*, Zichen Zhang*, [Savya Khosla](#), Ryan Marten, Derek Hoiem, Aniruddha Kembhavi
Computer Vision and Pattern Recognition, 2024
6. **Understanding and Improving Neural Active Learning on Heteroskedastic Distributions** [Link](#)
[Savya Khosla](#), Chew Kin Whye, Jordan T. Ash, Cyril Zhang, Kenji Kawaguchi, Alex Lamb
European Conference on Artificial Intelligence, 2023
7. **Interpolated Adversarial Training: Achieving Robust Neural Networks Without Sacrificing Too Much Accuracy** [Link](#)
Alex Lamb, Vikas Verma, Kenji Kawaguchi, Alexander Matyasko, [Savya Khosla](#), Juho Kannala, Yoshua Bengio
Neural Networks, 2022
8. **S-DCNN: Stacked Deep Convolutional Neural Networks for Malware Classification** [Link](#)
Anil Singh Parihar, Shashank Kumar, [Savya Khosla](#)
Multimedia Tools and Applications, 2022
9. **Catastrophic Failures of Neural Active Learning on Heteroskedastic Distributions** [Link](#)
[Savya Khosla](#), Alex Lamb, Jordan Ash, Cyril Zhang, Kenji Kawaguchi
NeurIPS 2021 Workshop on Distribution Shifts: Connecting Methods and Applications, 2021
10. **AE-DCNN: Autoencoder Enhanced Deep Convolutional Neural Network For Malware Classification** [Link](#)
Shashank Kumar*, [Savya Khosla](#)*, Shivangi Meena, Anil Singh Parihar
International Conference on Intelligent Technologies, 2021
11. **MuRIL: Multilingual Representations for Indian Languages** [Link](#)
Simran Khanuja, Diksha Bansal†, Sarvesh Mehtani†, [Savya Khosla](#)†, Atreyee Dey, Balaji Gopalan, Dilip Kumar Margam, Pooja Aggarwal, Rajiv Teja Nagipogu, Shachi Dave, Shruti Gupta, Subhash Chandra Bose Gali, Vish Subramanian, Partha Talukdar
ArXiv:2103.10730, 2021
Media Coverage: [Economic Times](#), [Indian Express](#), [Google AI Blog](#)

INDUSTRY EXPERIENCE

Google

Software Engineer

Bangalore, KA

Aug 2021 - Mar 2022

- Improved Google Search's web ranking infrastructure using deep learning for better multimodal document understanding
- Enhanced precision and recall in salient entity extraction from webpages by transitioning from traditional ML methods to LLMs

Google

Software Engineering Intern

Bangalore, KA

May 2020 - Jul 2020

- Initiated the development of MuRIL, a BERT-based multilingual language model for 17 Indian dialects
- Achieved 10.42% F1 improvement in sentiment analysis and 9.87% in named entity recognition for Indian languages

TEACHING EXPERIENCE

CS 445: Computational Photography

Teaching Assistant (Instructor: Prof. Derek Hoiem)

Urbana, IL

Aug 2023 - Dec 2023

CS 225: Data Structures and Algorithms with C++

Teaching Assistant (Instructor: Prof. Carl Evans and Prof. Brad Solomon)

Urbana, IL

Aug 2022 - May 2023

COURSES & CERTIFICATIONS

Graduate Courses

- CS 598: Machine Learning Algorithms for LLMs by Prof. Tong Zhang
- CS 598: Vision by Prof. Svetlana Lazebnik
- CS 588: Autonomous Vehicle System Engineering by Prof. David Forsyth
- CS 543: Computer Vision by Prof. Svetlana Lazebnik
- CS 445: Computational Photography by Prof. Derek Hoiem
- CS 444: Deep Learning for Computer Vision by Prof. Svetlana Lazebnik
- CS 410: Text Information System by Prof. ChengXiang Zhai
- CS 491: Competitive Programming by Prof. Mattox Beckman

Undergraduate Courses

- CO 201: Data Structures
- CO 202: Database Management System
- CO 203: Object-Oriented Programming
- CO 206: Algorithm Design and Analysis
- CO 304: Artificial Intelligence
- CO 404: Data Warehousing and Data Mining
- CO 407: Distributed Systems
- CO 414: Big Data Analytics
- CO 423: Swarm and Evolutionary Computing
- IT 420: Computer Vision

Online Courses & Certifications

- Deep Learning Specialization by Andrew Ng
- CS229 Lectures by Andrew Ng
- Introduction to TensorFlow for Artificial Intelligence, Machine Learning, and Deep Learning by deeplearning.ai
- Competitive Programming Bootcamp by Coding Blocks
- Machine Learning Master Course by Coding Blocks

PROFESSIONAL SERVICES

Organizer: [UIUC Vision Lunch](#) (Aug 2025 - Present), [Vision Mini-Conference](#) (2025)

Reviewer: TMLR

Mentor: [Undergraduate Research in Scientific Advancement](#) (Fall 2025)