

SURAJ KOTHAWADE

suraj.kothawade@utdallas.edu

Website

+1 213 285 2986

EDUCATION

- University of Texas at Dallas** (Aug'19 - Aug'23)
PhD, Computer Engineering GPA: 3.86/4.0
Advisor: Prof. Rishabh Iyer
- University of Southern California** (Jan'19 - Jul'19)
MS, Computer Science, GPA: 3.7/4.0
- SGGS Institute of Engineering & Technology** (July'14 - June'18)
B.Tech, Computer Science & Engineering, GPA: 8.93/10.0

RESEARCH INTERESTS

My research revolves around targeted data subset selection for improving the performance of machine learning models in realistic dataset scenarios like class imbalance, redundancy and out-of-distribution data. I'm interested in applying my research to a broad spectrum of domains in computer vision, natural language processing, speech, and recommendation systems.

INTERNSHIPS/WORK EXPERIENCE

- Google, Mountain View, CA** (May'22 - Aug'22)
Research Intern
 - Studied the effect of Global and Local attention in Vision Transformers
 - Designed a search space for Neural Architecture Search in order to automatically find the optimal backbone for any given downstream vision task.
- NVIDIA, Santa Clara, CA** (May'21 - Aug'21)
AI Research Intern
 - Designed a framework for auto-labelling traffic sign objects in images from multiple countries by using only few samples of a completely unknown traffic sign.
- NVIDIA, Santa Clara, CA** (May'20 - Jan'21)
AI Research Intern
 - Devised and implemented an algorithm for semantic object based retrieval called Deep Template Matching (DTM) that efficiently mines for semantically similar images based on a region of interest in a query image. (US Patent filed)
 - Demonstrated the efficacy of DTM for retrieving small and underrepresented objects.
 - Applied DTM for fixing false negative failure cases of an objected detector deployed in NVIDIA autonomous vehicles — by mining semantically similar objects from large unlabeled dataset and adding such images to training data.
- University of Texas at Dallas** (Aug'19 - May'20, Jan'21 - May'21, Aug'21 - May'22)
Research/Teaching Assistant (Fall 2021, Spring 2021: CS 6375,4375 - ML, Spring 2020, Fall 2019: CS 4348 - OS)
- University of Southern California, Los Angeles, CA** (Jan'19 - Aug'19)
Research Student (Advisor: Prof. Stefanos Nikolaidis)
 - Robotic Lime Picking using Cost-Based APF-RRTs by Modelling Leaves as Penetrable Obstacles
 - Learning Collaborative Action Plans from YouTube Videos — [ISRR 2019](#)
- Indian Institute of Technology, Bombay** (Dec'17 - Dec'18)
Research Intern (Advisor: Prof. Ganesh Ramakrishnan)
 - Worked with on developing machine learning models to solve computer vision problems in CCTV videos.
 - Lead a team to deliver a Compliance and Quality Monitoring System for the **Ministry of Rural Development**: for the following compliances: 1) Predict if a Class has Started or Not (Implemented Handcrafted features for higher accuracy). 2) Classroom Attendance and 3) Uniform Detection. (Used multi-class customized YOLOv2 on edge devices). [video](#)
 - Developed an open source toolkit for Visual Data Subset Selection and Summarization: [arXiv-preprint](#) [GitHub](#)
- Aitoe Labs** (Apr'18 - Dec'18)
Machine Learning Engineer
 - Implemented machine learning pipeline and system architecture for analyzing big data from 500+ CCTV cameras in Bhopal(IN) to deliver person search, face search, face recognition and text search for the state police department in Madhya Pradesh.
 - Solved problems like scheduling tasks efficiently to GPUs, storage and retrieval of huge metadata for quick search.

RESEARCH PAPERS

PREPRINTS

- [1] Suraj Kothawade, Shivang Chopra, Saikat Ghosh, and Rishabh Iyer. **Active Data Discovery: Mining Unknown Data using Submodular Information Measures**. In *arXiv preprint arXiv:2206.08566*.
- [2] Suraj Kothawade, Saikat Ghosh, Sumit Shekhar, Yu Xiang, and Rishabh Iyer. **TALISMAN: Targeted Active Learning for Object Detection with Rare Classes and Slices using Submodular Mutual Information**. In *arXiv preprint arXiv:2112.00166*.

- [3] Suraj Kothawade, and Rishabh Iyer. **PROBE: Deep Submodular Networks for Subset Selection**. In *arXiv preprint arXiv:2010.08593*.

CONFERENCE PAPERS

- [1] Changbin Li*, Suraj Kothawade*, Feng Chen, Rishabh Iyer. **PLATINUM: Semi-Supervised Model Agnostic Meta-Learning using Submodular Mutual Information**. To appear at *The Thirty-ninth International Conference on Machine Learning*, ICML 2022.
- [2] Suraj Kothawade, Vishal Kaushal, Ganesh Ramakrishnan, Jeff Bilmes, and Rishabh Iyer. **PRISM: A Unified Framework of Parameterized Submodular Information Measures for Targeted Data Subset Selection and Summarization**. In *Thirty-Sixth AAAI Conference on Artificial Intelligence*, AAAI 2022.
- [3] Suraj Kothawade, Donna Roy, Michele Fenzi, Elmar Haussman, Jose M. Alvarez, and Christoph Angerer. **Object Level Targeted Selection using Deep Template Matching**, In *33rd IEEE Intelligent Vehicles Symposium, IV 2022*
- [4] Suraj Kothawade, Nathan Beck, Krishnateja Killamsetty and Rishabh Iyer, **SIMILAR: Submodular Information Measures Based Active Learning In Realistic Scenarios**, In *35th Conference on Neural Information Processing Systems, NeurIPS 2021*.
- [5] Vishal Kaushal, Rishabh Iyer, Suraj Kothawade, Rohan Mahadev, Khoshrav Doctor and Ganesh Ramakrishnan, **Learning from less data: A unified data subset selection and active learning framework for computer vision**, In *2019 IEEE Winter Conference on Applications of Computer Vision (WACV)*, pp. 1289-1299, 2019.
- [6] Vishal Kaushal, Sandeep Subramanian, Suraj Kothawade, Rishabh Iyer and Ganesh Ramakrishnan, **A framework towards domain specific video summarization**. In *2019 IEEE Winter Conference on Applications of Computer Vision (WACV)*, pp.666-675, 2019.
- [7] Vishal Kaushal, Rishabh Iyer, Khoshrav Doctor, Anurag Sahoo, Pratik Dubal, Suraj Kothawade, Rohan Mahadev, Kunal Dargan and Ganesh Ramakrishnan, **Demystifying multi-faceted video summarization: Tradeoff between diversity, representation, coverage and importance**. In *2019 IEEE Winter Conference on Applications of Computer Vision (WACV)*, pp. 452-461, 2019.
- [8] Heramb Nemlekar, Ziang Liu, Suraj Kothawade, Sherdil Niyaz, Barath Raghavan, Stefanos Nikolaidis, **Robotic Lime Picking by Considering Leaves as Permeable Obstacles**, In *International Conference on Intelligent Robots and Systems (IROS 2021)*
- [9] Hejia Zhang, Po-Jen Lai, Sayan Paul, Suraj Kothawade and Stefanos Nikolaidis. **Learning collaborative action plans from youtube videos**. In *Proceedings of the International Symposium on Robotics Research (ISRR 2019)*, Hanoi, Vietnam, 2019.

WORKSHOP PAPERS

- [1] Suraj Kothawade, Shivang Chopra, Saikat Ghosh, and Rishabh Iyer. **Active Data Discovery: Mining Unknown Data using Submodular Information Measures**. In *ReALML, ICML 2022, Adaptive Experimental Design and Active Learning in the Real World*.
- [2] Suraj Kothawade, Donna Roy, Michele Fenzi, Elmar Haussman, Jose M. Alvarez, and Christoph Angerer. **Object Level Targeted Selection using Deep Template Matching**, *Spotlight Talk at Machine Learning for Autonomous Driving Workshop at the 35th Conference on Neural Information Processing Systems (NeurIPS 2021)*
- [3] Suraj Kothawade, Lakshman Tamil, Rishabh Iyer. **Targeted Active Learning using Submodular Mutual Information for Imbalanced Medical Image Classification**, *Medical Imaging Meets NeurIPS Workshop at the 35th Conference on Neural Information Processing Systems (NeurIPS 2021)*
- [4] Suraj Kothawade, Vishal Kaushal, Ganesh Ramakrishnan, Jeff Bilmes, and Rishabh Iyer. **Submodular Mutual Information for Targeted Data Subset Selection**, In *ICLR 2021 Workshop: From Shallow to Deep: Overcoming Limited and Adverse Data*.
- [5] Suraj Kothawade, Nathan Beck, Krishnateja Killamsetty and Rishabh Iyer, **SIMILAR: Submodular Information Measures Based Active Learning In Realistic Scenarios**, In *ICML 2021 Workshop: Subset Selection in Machine Learning*.
- [6] Suraj Kothawade, Vinaya Khandelwal, Huaduo Wang, Kinjal Basu, Gopal Gupta. **AUTO-DISCERN: Autonomous Driving Using Common Sense Reasoning**, In *ICLP 2021 Workshop on Goal-directed Execution of Answer Set Programs*.
- [7] Vishal Kaushal, Suraj Kothawade, Rishabh Iyer and Ganesh Ramakrishnan. **Realistic Video Summarization through VI-SIOCIETY: A New Benchmark and Evaluation Framework**. In *Proceedings of the ACM MM 2nd International Workshop on AI for Smart TV Content Production, Access and Delivery*, pp. 37-44, 2020.

DEMONSTRATIONS

- [1] Suraj Kothawade and Rishabh Iyer. **An Efficient Data Exploration Framework for Effective Learning**. At *33rd IEEE Intelligent Vehicles Symposium, IV 2022*.

PATENTS

1. US Patent Application 17/226584, "Scalable Semantic Image Retrieval In The Wild With Deep Template Matching," Apr 9, 2021. Inventors: Donna Roy, Suraj Kothawade, Michele Fenzi, Elmar Haussman, Jose M. Alvarez, and Christoph Angerer.
2. US Patent Application 17/689799, "Object Data Curation of Map Information Using Neural Networks For Autonomous Systems and Applications," March 15, 2022. Inventors: Michele Fenzi, Suraj Kothawade, Nisan Haramati, Ozan Tonkal, Christoph Angerer.

ACHIEVEMENTS & AWARDS

- *Jan Van der Ziel Fellowship* - University of Texas at Dallas.
- *Runner-up* at the University of Texas at Dallas 2022 Three Minute Thesis (3MT) Competition.
- ICML 2022 Participation Grant.
- Accepted for Doctoral Consortium at WACV 2022.
- **Best Student Award** 2018 by Tata Sons. (Awarded to 1 out of 630 students across all engineering departments)
- **Best Project Award** 2018 by Tata Sons. (Awarded to 1 out of 30 projects across all engineering departments)
- Only student to receive both, the Best Student Award and Best Project Award awards since 1981 by Tata Sons
- **ACM ICPC 2017** Honorable mention.
- Ranked **1st/160** in Computer science & Engineering department in Junior year also cumulatively ranked **2nd/160**.

PROFESSIONAL ACTIVITIES/SERVICE

- Program Committee Member for NeurIPS 2022 Main Track, NeurIPS 2022 Datasets and Benchmarks Track, ICML 2022, ICLR 2022, WACV 2022, AAAI (2021,2022), SIGKDD 2022, CVPR 2021, IROS 2021

SOFTWARE CONTRIBUTIONS

- **DISTIL: Deep dIversified inTeraCTive Learning** (<https://github.com/decile-team/distil>)
DISTIL implements a number of state-of-the-art active learning algorithms.
- **TRUST: TaRgeted sUbSet selecTion** (<https://github.com/decile-team/trust>)
TRUST supports a number of algorithms for targeted selection which provides a mechanism to include additional information via data to prioritize the semantics of the selection.

PROGRAMMING SKILLS

- **Languages:** Python, C++, C
- **Frameworks:** PyTorch, Tensorflow, Caffe, Keras, OpenCV