Harsh Maheshwari

✉ harshm121@gmail.com harshm121.github.io +91-9654011364

EDUCATION

Georgia Institute of Technology, *Atlanta, US*
Master of Science (Thesis) in Computer Science (Specialization: ***Machine Learning****); GPA: 4.0/4.0* *Aug. 2021 – May. 2023*
Advisor: ***Prof. Devi Parikh, Prof. Zsolt Kira****; Thesis topic: Semi-supervised Semantic Segmentation*

Indian Institute of Technology, Delhi, *New Delhi, India*
B.Tech in Electrical Engineering; ***Advisor:*** ***Prof. Prathosh AP****; GPA: 8.27/10.0* *Jul. 2015 – May. 2019*

PUBLICATIONS (relevant)

1. **Missing Modality Robustness in Semi-Supervised Multi-Modal Semantic Segmentation** ,
[***WACV'24***] *Authors: Harsh Maheshwari, Yen-Cheng Liu, Zsolt Kira*
2. **We're Not Using Videos Effectively: An Updated Domain Adaptive Video Segmentation Baseline**,
[***TMLR***] *Authors: Simar Kareer, Vivek Vijaykumar, Harsh Maheshwari, Prithvijit Chattopadhyay, Judy Hoffman, Viraj Prabhu*
3. **Recommendation of Compatible Outfits Conditioned on Style** ,
[***ECIR'22 (oral)***] *Authors: Harsh Maheshwari*, Lucky Dhakad*, Debopriyo Banerjee*, Niloy Ganguly, Arnab Bhattacharya*
4. **CoSIR: Optimal control of SIR epidemic dynamics by mapping to Lotka-Volterra System**,
[***ICLR'21 Workshop MLPCP***] *and* [***CHIL'21 Workshop***]
Harsh Maheshwari, Shreyas Shetty, Nayana Bannur, Srujana Merugu

WORK EXPERIENCE

Research Engineer III - Avataar, *Bengaluru*
Generative AI; Vision. *RE III - Feb. 2025 – Present.*
Managers: Dr. Shubham Goel, Dr. Sohil Shah *RE II - July. 2023 – Jan. 2025*

- Novel View Synthesis:
  - Worked on the ambitious problem statement of **sparse** (3-4) **images to Novel View Synthesis** of an object, learning to rely on available information and hallucinate missing information whenever needed. **[Presentation]**
  - Fine-tuned a **stable diffusion** model (Zero123-like) that is conditioned on sparse images by devising novel ways of conditioning diffusion model and extensive error analysis of baselines.
  - Brought down training time from 12 days (On 8xA100-80GB) to 2 days following tricks from the literature.
  - Handled large-scale training (∼1B parameters, 32xH100s) and maintained the cluster for efficient use (4x8xH100).
  - Put the model into production for demo to potential clients for various e-commerce use cases. **[Demo]**

- Lifestyle Image generation:
  - Led the work on generating Lifestyle Images with the object image as input and optional human edits. **[Eg. outputs]**
  - Designed an automated system combining LLMs and Flux models to generate stylized prompts, place the object on a canvas, and create cohesive lifestyle imagery. **[Demo]**

- *Other Misc. work*:
  - Devised a relighting algorithm to make 2.2D videos for e-commerce objects. **[Sample]**
  - Improved distillation of the learnt 3D prior using Signed Density Fields to reconstruct 3D object.
  - Took interviews and played an active role in hiring for research and applied ML roles.

Applied Scientist (internship) - Amazon, *Seattle, WA*
Personalizations team in Amazon Fashion *May. 2022 – Aug. 2022*

- Formulated and proposed the use of **active learning** for exploration in recommendations to drive customer understanding.
- Improved ROCAUC score for predicting the like-rate of the recommendations by **4% points with limited data per customer.**
- Devised an offline evaluation framework to measure the goodness and iterate over multiple recommendation policies.

Return offer: *Was extended a* ***full-time Applied Scientist*** *role ("inclined-to-hire") for the Fashion Personalization team*

Data Scientist II - Flipkart, *Bengaluru, India*
Largest E-Commerce platform in India with over 200M users *July. 2019 – July. 2021*

- *Complete The Look* (Prof. Niloy Ganguly - IIT KGP, Dr. Arnab Bhattacharya - Flipkart):
  - Problem: Generating **fashion-compatible** and **diverse** outfits for a 'hero' product for Indian users and their preferences.
  - Designed an architecture and algorithm to learn outfit compatibility conditioned under a 'style'.
  - Implemented SOTA **fashion-compatibility, apparel segmentation & classification** models and a flask tool for annotations.
  - *Publication:* Recommendation of Compatible Outfits Conditioned on Style, ECIR'22 (Oral)
- *Candidate Generation and Ranking* (Samik Datta, Dr. Adiya Rachakonda - Flipkart):
  - Customized **Bayesian Personalised Ranking based Matrix Factorisation** framework for Fkart's homepage recommendation
  - *Impact:* Improvement in conversion by 2bps (units/visits) and 16bps (units/visitor)
  - *Impact:* Won an internal **project award** for this work

Covid19 Volunteer - DSIndiaVsCovid19, Wadhwani AI,
A consortium of volunteer technologists to support public authorities in managing COVID-19 *March. 2020 – July. 2021*

- *Forecasting* (Dr. Srujana Merugu, Dr. Alpan Raval, Dr. Mohit Kumar):
  - Developed an **ML framework** for infectious disease forecasting based on **SEIR epidemiological model variants**
  - *Impact:* The system is being used for COVID-19 medical preparedness in war rooms of heavily impacted Indian cities.

Projects

Multi-modal Semi Supervised Semantic Segmentation, *Advised by: Prof. Zsolt Kira*
MS Thesis *Jan. 2022 – May. 2023*

- Led work on **semi-supervised** (label-efficient) semantic **segmentation** using multiple spatial modalities (RGB, Depth).
- Reproduced results on several state-of-the-art well well-cited baselines in semi-supervised semantic segmentation and modality fusion techniques (using CNNs and **Transformers**)
- Created a novel multi-modal algorithm for effectively using unlabeled data for semantic segmentation while making the model robust to missing modalities at test time.
  - *Publication:* Missing Modality Robustness in Semi-supervised Multi-modal Semantic Segmentation, WACV'24

Misc course projects,
Done for various courses during MS at Georgia Tech *Aug. 2021 – May. 2023*

- Generating *consistent* images using text prompts by incorporating CLIP in VQ-GAN's latent space. (with Prof. Devi Parikh)
- Visual Question Answering using CLIP: Using CLIP to solve the VQA task in a zero-shot/few-shot setting (Project Page)
- Unsupervised Domain Adaptation: Used FixMatch consistency to achieve 4% improvement over the state-of-the-art approach for Unsupervised Domain Adaptation from SVHN to MNIST.

SCHOLASTIC ACHIEVEMENTS

- Earned a **Research Assistantship** for all semesters of M.S. at Georgia Tech, funded by Prof. Devi Parikh (*2021-23*)
- Awarded **Best Project Award** at Flipkart for a recommendations ranking project (*2020*)
- Among 11 finalist teams in 4-stage National level AI/ML Challenge - Flipkart GRiD (*2019*)
- **Huawei:** Among **4 students** from India selected for a 2-week training program in Huawei Headquarters, Shenzhen (*2018*)
- **All India Rank 834** in IIT-Joint Entrance Exam Advanced among **1.4 million** students. (*2015*)
- **NSEP top 1%: Top 1%** out of 37837 in National Standard Examination in Physics organised by IAPT (*2015*)

SKILLS & INTERESTS

Research Interest: Computer Vision, Diffusion Models, Creative AI, Vision + Language

Deep Learning Frameworks: PyTorch, TensorFlow, Keras

Languages: Python, C++, Java, Hive, SQL

Courses: ML with Limited Supervision, Deep Learning, Big Data Systems, Mathematical Foundations of ML, Introduction to Machine Learning, Advanced Machine Learning, Computational Learning Theory and Mind, Information bottleneck Theory of Deep Learning, Information Theory, Data Structures and Algorithms, Probability, Linear Algebra