Robik Shrestha

robikshrestha@gmail.com • +1 (585) 766-5549 • LinkedIn • Google Scholar • Github

Summary

- 6+ years of research and engineering experience with deep learning systems
- 9+ publications on improving the robustness of deep learning systems. Published at: CVPR, NeurIPS, ACL, ECCV (oral), Frontiers etc
- 3 industry collaborations during my Ph.D. on large-scale foundation models
- Founding member of Viveka Health, a successful startup dealing with fraud in US healthcare claims

Education

Rochester Institute of Technology — ROCHESTER, NY
Ph.D. (Thesis: Towards Bias-Resilient Deep Neural Networks) — *Aug '17 – Dec '23*
Highlights:
- 3+ publications showcasing shortcomings of "state-of-the-art" models, including foundation models: right for the wrong reasons, failure on higher-level cognitive tests and realistic distribution shifts
- 4+ publications improving robustness and efficiency of vision and language systems, while providing greater simplicity compared to the existing methods

Institute of Engineering, Tribhuvan University — LALITPUR, NEPAL
B.E. in Computer Engineering — *Oct '08 – Dec '12*

Experience

Amazon AGI — SEATTLE, WA
Incoming Applied Scientist — *May '24 – Future*
- Will be working on developing novel architectures and methodologies for state-of-the-art multimodal systems

University of Rochester — ROCHESTER, NY
Lead Research Engineer (Temporary Role) — *Feb '24 – Apr '24*
- Developed a prototype system for generative AI to optimize inertial confinement fusion simulating multivariate, temporal observations.
- Also collaborated with Dr. Christopher Kanan's Lab projects related to multimodal LLMs.

Amazon AWS AI Labs — BELLEVUE, WA
Applied Research Scientist Intern — *May '23 – Nov '23*
- Built FairRAG, a retrieval augmented generation framework to foster fairness and diversity in generative AI (e.g., stable diffusion) (CVPR'24)
- Devised efficient conditioning mechanism and simple, posthoc debiasing strategies to enhance fairness.
- Also worked on red teaming efforts to prevent toxic/harmful content generation and IP infringements

SRI International — ROCHESTER, NY (REMOTE)
Research Assistant (funded my Ph.D.) — *Apr '20 – Jan '23*
- Constructed a comprehensive evaluation setup to test cognitive abilities of VQA models using Bloom's taxonomy of cognition
- Identified deficiencies in understanding, analysis, reasoning, and generative skills, providing valuable insights for model enhancement (Under Review)

Adobe — ROCHESTER, NY (REMOTE)
Research Intern — *May '21 – Oct '21*
- Formulated a novel contrastive learning algorithm, enhancing text-to-image retrieval skills
- Elevated sensitivity in multimodal embeddings to discern fine-grained conceptual differences

ITEL Labs — ROCHESTER, NY (REMOTE)
Research Assistant (funded my Ph.D.) — *Jul '18 – Mar '21*
- Overhauled a manual system, transitioning it to a semi-automated framework and automating 85% of inputs for operational efficiency
- Engineered ML systems (edge and server-side) handling damaged flooring data in insurance claims
- Significantly improved bias resilience, calibration, and OOD detection for robust ML deployment

Viveka Health LALITPUR, NEPAL

Full Stack Developer *Feb '14 – Apr '16*

- Saved $1M+ for 25K+ lives by tackling fraud, waste, and abuse in the U.S. healthcare claims
- Solo-built full-stack proof-of-concept: ETL, database, UI, fraud detection
- Led the role definition for a 10+ member team, ensuring project success and operational efficiency, contributing to overall effectiveness

Skills

- 6+ years of programming experience with Python utilizing deep/machine learning frameworks: PyTorch, Lightning, HuggingFaces, SciPy, Pandas

- Hands-on Experience with large-scale deep learning models including: CLIP, BLIP, BERT, Claude, GPT-3.5/GPT-4V, LLaVa, Stable Diffusion

- Experience building web services with Python, Javascript, MongoDB (NoSQL), Redis

Selected Publications and Pending Patent(s)

- "FairRAG: Fair Retrieval Augmentation for Fair Generation." Robik Shrestha, Yang Zou, James Chen et al (Work done at Amazon, '23) (CVPR'24)

- "BloomVQA: Assessing Hierarchical Multi-modal Comprehension." Yunye Gong*, Robik Shrestha*, Jared Claypoole et al (* = equal contributions) (Work done at SRI, '23) (Under Review)

- "OccamNets: Mitigating Dataset Bias by Favoring Simpler Hypotheses." Robik Shrestha, Kushal Kafle, Christopher Kanan (ECCV '22) (Oral, Top 2.7%)

- "On the value of out-of-distribution testing: An example of Goodhart's law." Damien Teney, Kushal Kafle, Robik Shrestha et al (NeurIPS '20)

- "REMIND Your Neural Network to Prevent Catastrophic Forgetting." Tyler Hayes*, Kushal Kafle*, Robik Shrestha*, Manoj Acharya, and Christopher Kanan (ECCV '20) (* = equal contributions)

- "A Negative Case Analysis of Visual Grounding Methods for VQA." Robik Shrestha, Kushal Kafle, and Christopher Kanan (ACL '20)

- "Answer Them All! Toward Universal Visual Question Answering Models." Robik Shrestha, Kushal Kafle, and Christopher Kanan (CVPR '19)

- **Pending Patent:** C. Kanan, T.L. Hayes, K. Kafle, and R. Shrestha. "Method for Training Parametric Machine Learning Systems," January 28, 2021. US Patent App. 16/938,035

- Visit my Google Scholar Page for the full list!

Invited Talks and Guest Lectures

- "OccamNets: Mitigating Dataset Bias by Favoring Simpler Hypotheses" [Oral Presentation], European Conference on Computer Vision (ECCV) '22, Tel Aviv, Israel

- "Dataset Bias in Vision and Language Tasks: Problems and Potential Solutions" [Invited Talk], Center for Human Aware AI (CHAI) '21, Rochester, NY

- "Dataset Bias and Bias Mitigation Techniques" [Guest Lecture], Deep Learning for Vision '20/21, RIT

- "Answer Them All! Toward Universal Visual Question Answering Models" [Spotlight], Language and Vision Workshop, CVPR '19, Long Beach, California