Arijit Ray

array@bu.edu https://arijitray.com Google Scholar

Goal

To build machines that can *perceive* the world, *interact* with humans, and *think*- to solve tasks in the digital and physical worlds.

Appointments

Jun'24 - ?	**Student Collaborator**, Allen Institute for AI, PRIOR Team *Spatial reasoning in multimodal language models, Multimodal browser-use agents.*
May'25 - Dec'25	**Research Intern/Student Researcher**, Google *Working with Pixel and Deepmind Foundational Research (Leonidas Guibas).*
Summer'23	**AI Resident**, Google X Moonshot Labs *Mineral Team- adapting multimodal language models for custom phrase localization in images.*
Summer'22	**Research Scientist Intern**, Meta (Facebook) AI (FAIR-Accel) *Explored adaptation strategies for compositional reasoning in vision-language models.*
2017 - 2021	**Computer Scientist**, SRI International *Lead contributor on the SRI-UToronto Team for the DARPA Explainable AI Program.*
Summer'16	**Deep Learning Intern**, Blue River Technology *Developed plant localization- key selling point leading to John Deere acquisition for* $\$305M$.
2016 - 2017	**Graduate Research Assistant**, Virginia Tech *Improving models that can answer questions about images. Advised by Devi Parikh.*

Education

2021 – ?	**Ph.D., Boston University**, Computer Science *Teaching machines to reason about the 3D visual world through language.* *Advised by Kate Saenko, Bryan A. Plummer, and Ranjay Krishna (**University of Washington**)*
2015 – 2017	**M.S., Virginia Polytechnic Institute and State University**, Computer Engineering Thesis: *Developing models that can converse with humans, advised by Devi Parikh.*
2011 – 2015	**B.Tech., SRM University, India**, Electrical Engineering *GPA:* 9.05/10*, First-Class Distinction. Received Academic Merit Scholarship.*

Research Publications

Working/Pre-prints

1. Ellis II Brown, **Arijit Ray**, Ranjay Krishna, Ross Girshick, Rob Fergus, and Saining Xie. *SIMs: Can Simple Simulators Elicit Spatial Understanding in Video Language Models?* preprint. 2025.
2. **Arijit Ray**, Ahmed Abdelkader, Chengzhi Mao, Bryan A. Plummer, Kate Saenko, Ranjay Krishna, Leonidas Guibas, and Wen-Sheng Chu. *Mull-tokens: Modality-Agnostic Latent Thinking*. preprint. 2025.

Peer-reviewed Conferences

1. Abhay Deshpande, Yuquan Deng, **Arijit Ray**, Jordi Salvador, Winson Han, Jiafei Duan, Kuo-Hao Zeng, Yuke Zhu, Ranjay Krishna, and Rose Hendrix. "GraspMolmo: Generalizable Task-Oriented Grasping via Large-Scale Synthetic Data Generation". In: *Conference on Robot Learning (CoRL)* (2025).

2. **Arijit Ray**, Jiafei Duan, Ellis Brown, Reuben Tan, Dina Bashkirova, Rose Hendrix, Kiana Ehsani, Aniruddha Kembhavi, Bryan A. Plummer, Ranjay Krishna, Kuo-Hao Zeng, and Kate Saenko. "SAT: Dynamic Spatial Aptitude Training for Multimodal Language Models". In: *Conference on Language Modeling (COLM)* (2025).

3. Jimuyang Zhang, Zanming Huang, **Arijit Ray**, and Eshed Ohn-Bar. "Feedback-Guided Autonomous Driving". In: *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)* (2024). (Highlight, top 2.8%).

4. **Arijit Ray**, Filip Radenovic, Abhimanyu Dubey, Bryan A Plummer, Ranjay Krishna, and Kate Saenko. "COLA: A Benchmark for Compositional Text-to-image Retrieval". In: *Conference on Neural Information Processing Systems (NeurIPS)* (2023).

5. Reuben Tan, **Arijit Ray**, Andrea Burns, Bryan A Plummer, Justin Salamon, Oriol Nieto, Bryan Russell, and Kate Saenko. "Language-Guided Audio-Visual Source Separation via Trimodal Consistency". In: *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)* (2023), pp. 10575–10584.

6. Kamran Alipour, **Arijit Ray**, Xiao Lin, Michael Cogswell, Jurgen P Schulze, Yi Yao, and Giedrius T Burachas. "Improving users' mental model with attention-directed counterfactual edits". In: *Applied AI Letters* 2.4 (2021), e47.

7. **Arijit Ray**, Michael Cogswell, Xiao Lin, Kamran Alipour, Ajay Divakaran, Yi Yao, and Giedrius Burachas. "Generating and evaluating explanations of attended and error-inducing input regions for VQA models". In: *Applied AI Letters* 2.4 (2021), e51.

8. Kamran Alipour, **Arijit Ray**, Xiao Lin, Jurgen P Schulze, Yi Yao, and Giedrius T Burachas. "The impact of explanations on AI competency prediction in VQA". In: *IEEE International Conference on Humanized Computing and Communication with Artificial Intelligence (HCCAI)* (2020), pp. 25–32.

9. **Arijit Ray**, Karan Sikka, Ajay Divakaran, Stefan Lee, and Giedrius Burachas. "Sunny and Dark Outside?! Improving Answer Consistency in VQA through Entailed Question Generation". In: *Conference on Empirical Methods in Natural Language Processing (EMNLP)* (2019), pp. 5860–5865.

10. **Arijit Ray**, Yi Yao, Rakesh Kumar, Ajay Divakaran, and Giedrius Burachas. "Can you explain that? Lucid explanations help human-AI collaborative image retrieval". In: *AAAI Conference on Human Computation and Crowdsourcing* 7.1 (2019), pp. 153–161.

11. **Arijit Ray**, Gordon Christie, Mohit Bansal, Dhruv Batra, and Devi Parikh. "Question relevance in VQA: identifying non-visual and false-premise questions". In: *Conference on Empirical Methods in Natural Language Processing (EMNLP)* (2016).

Workshops/Non-archival

1. **Arijit Ray**, Dina Bashkirova, Reuben Tan, Kuo-Hao Zeng, Bryan A Plummer, Ranjay Krishna, and Kate Saenko. *R2D3: Imparting Spatial Reasoning by Reconstructing 3D Scenes from 2D Images*. preprint. 2024.

2. Dina Bashkirova, **Arijit Ray**, Rupayan Mallick, Sarah Bargal, Jianming Zhang, Ranjay Krishna, and Kate Saenko. *Lasagna: Layered Score Distillation for Disentangled Object Relighting*. arxiv:2312.00833. 2023.

3. Katherine Deng, **Arijit Ray**, Reuben Tan, Saadia Gabriel, Bryan Plummer, and Kate Saenko. *Socratis: Are large multimodal models emotionally aware?* ICCV Worskshop on Emotionally and Culturally Aware AI (Oral). 2023.

4. Shalini Ghosh, Giedrius Burachas, **Arijit Ray**, and Avi Ziskind. *Generating Natural Language Explanations for Visual Question Answering Using Scene Graphs and Visual Attention*. IJCAI/ECAI Workshop on Explainable Artificial Intelligence, XAI 2018. 2018.

Patents

1. Giedrius Burachas, **Arijit Ray**, and Yi Yao. *Attention-based explanations for artificial intelligence behavior*. US Patent 10,909,401. Feb. 2021.
2. Ajay Divakaran, Karan Sikka, **Arijit Ray**, Xiao Lin, and Yi Yao. *User targeted content generation using multimodal embeddings*. US Patent App. 17/191,698. Sept. 2021.

Awards

2019 **SRI Shark Tank Award**, SRI International, Center for Vision Technologies.
Received $50,000 *for* 6 *months that supported my project on generating personalized content to convey effects of climate change. Awarded to 3 projects in the center.*

2016 **Employee of the Fortnight**, Blue River Technology.
For training a state-of-the-art plant detection model, a key selling point for the acquisition by John Deere.

2013 **Silver Medal**, Research Day, SRM University.
Rank 2 *out of* ~300 *students in the department.*

2012 **Academic Merit Scholarship**, SRM University.
Rank 3 *out of* ~300 *students in the Electrical Engineering Department.*

Mentoring

2023 Gitika Jha (AI4All Undergraduate BU; now SDE at Amazon)
Katherine Deng (AI4All Undergraduate BU; now SWE at Fidelity Investments)
Jiayi Shen (AI4All Undergraduate BU; now MS student at Brown University)
Xavier Thomas (MS BU; Now PhD student at BU)

2022 Praneeth Chandra Bogineni (MS BU; now at a startup, Oplus.ai)

2018-2021 Kamran Alipour (UC San Diego; now Senior AI R&D Engineer, Williams Sonoma)

2019 Julia Kruk (SRI International; now at Meta AI)

Leadership/Venture

Summer'24 **Build @ Pillar VC**
Selected as one of the Build @ Pillar VC summer cohort for entrepreneurship networking and ideation in the robotics space.

Spring'23 **Student Leader**, AI For Impact Venture Studio, MIT
Part of the student leadership council organizing networking events with over 100 attendees from 5 schools in the Boston area.

2021 – 2022 **Co-chair**, AI+X of BU and Harvard
Started a graduate student workshop investigating how AI can impact contemporary research areas.

2016 – 2017 **Vice President**, Tau Beta Pi Engineering Honor Society
Vice President of the Virginia Tech Chapter

Professional Service

2016 – ? **Reviewer**
Neurips'22-24, CVPR'26,25,24,16, ECCV'24, ICLR'25, EMNLP'23, COLING'22.

2022, 2017 **Judge**, Blue Ridge Highlands Regional Science Fair
Science fair for high-school students

Media

2023 **The Generative AI Podcast**
I was featured in a podcast about using AI to predict social media responses.

2019 **TechXplore, Phys.org**
An image-guessing game to evaluate the helpfulness of machine explanations, presented also as a CVPR 2019 Demo and AAAI HCOMP 2019 Poster.

2014 **Indian Express, Deccan Chronicle, Engineering.Careers360**
Prototyped an Unmanned Autonomous Drone for identifying disaster victims.

Early Achievements

2011 All India Undergraduate Entrance Examination (SRM-JEE)
99%*ile among students in India.*

All India Central Board Examinations
99%*ile math score among students in India.*

2007 National Science Olympiad
All India Rank: 168*, City Rank:* 7*.*

2006 Founded middle-school science society
Goal of encouraging middle-school students to take an interest in science. Won accolades in multiple school/city-level exhibitions.