Shraman Pramanick

Curriculum Vitae

108 W 39th St, Apt 28
Baltimore, MD 21210
+1 (443) 722 4993
spraman3@jhu.edu
Personal Website

Research Interests

- **Self-Supervised Learning**, **Vision-Language Pre-training**
- **Multimodal Learning** (Vision + Language, Vision + Other Modalities)
- Planet-Scale **Image** & **Video Geo-localization** (Beyond Retrival)

Education

Jan 2021 - Present — **Johns Hopkins University**, *Baltimore, MD, USA.*
Ph.D. (with M.S.) in Electrical and Computer Engineering
Advisor: Rama Chellappa, AIEM Lab, ECE (**GPA:** 4.0/4.0)
Research: Self-Supervised Learning, Multimodal Learning, Image Geo-localization

2016 - 2020 — **Jadavpur University**, *Kolkata, WB, India.*
Bachelor of Engineering (B.E.) in Electronics & Telecommunication Engineering
Advisor: Amit Konar, AI Lab, ETCE (**GPA:** 9.41/10.0)
Bachelor's Thesis: Localizing and Grasping of 3-D Objects by a Vision-Actuated Robot Arm using Brain-Computer Interface

Research Experience

June 2023 - Present — **Research Scientist Intern**, *Meta AI.*
Collaborators: Nicolas Ballas, Amjad Almahairi, Guangxing Han, Rui Hou, & Qifan Wang
• **Visual Instruction Tuning:** Working on LLM-based framework for open-ended, customizable and unified fine- and coarse-grained vision-centric tasks.

May 2022 - Mar 2023 — **Research Scientist Intern**, *Meta AI.*
Collaborators: Pengchuan Zhang, Li Jing, Yale Song, Hardik Shah, & Yann LeCun
• **Egocentric Video-Language Pre-training:** Worked on egocentric video-language foundational model using cross-modal *fusion* in uni-modal backbones, and achieved state-of-the-art performance on EgoMCQ, EgoNLQ, EgoMQ, QFVS, EgoTaskQA, CharadesEgo, Epic-100.
• **Multimodal Dimension-Contrastive Pre-training:** Proposed VoLTA, a dimenison-contrastive pre-training for image-caption pairs with explicit region-level understanding.

Feb 2021 - Present — **Graduate Research Assistant**, *Johns Hopkins University.*
Advisor: Rama Chellappa, AIEM Lab, ECE
• **Planet-scale Single Image Geo-localization:** Proposed TransLocator, a dual-branch transformer network for planet-scale image geo-location as a part of WRIVA program.
• **Real-time Detection of Activities in Untrimmed Videos:** Working on proposal-based solution to spatio-temporal action detection in untrimmed videos as a part of DIVA program.

May 2020 - Jan 2021 — **Research Associate**, *QCRI (Doha) & IIIT-Delhi Collaboration.*
Advisor: Preslav Nakov & Tanmoy Chakraborty
• **Data Efficient and Scalable Multimodal Learning (Vision, Language & Speech)**
- Multimodal Abstractive Summarization, Detecting Harmful Internet Memes and Their Targets.

May 2019 - Aug 2019 — **Mitacs Globalink Research Intern**, *University of Montreal, Canada.*
Advisor: Antoine Saucier, Mathematical and Industrial Engineering
• Worked on classical NR algorithms that preserve details, edges and fine patterns in images.

Teaching Experience

Spring 2023 — **Coarse Assistant: Machine Intelligence (EN.520.650)**, *Johns Hopkins University.*
Spring 2022 — **Coarse Assistant: Machine Intelligence (EN.520.650)**, *Johns Hopkins University.*

Selected Publications

Please see Google Scholar, Semantic Scholar for the complete list of publications

Pre-prints

- **Pramanick S.**, Song Y., Nag S., Lin K., Shah H., Shou M., Chellappa R., Zhang P., "EgoVLPv2: Egocentric Video-Language Pre-training with Fusion in the Backbone". (Under Submission) [Project]
- **Pramanick S.***, Jing L.*, Nag S.*, Zhu J., Shah H., LeCun Y., Chellappa R., "VoLTA: Vision-Language Transformer with Weakly-Supervised Local-Feature Alignment". (Under Submission) [Paper]

Conference Proceedings

- **Pramanick S.**, Nowara E.M., Gleason J., Castillo C.D., Chellappa R., "Where in the World is this Image? Transformer-based Geo-localization in the Wild", *European Conference on Computer Vision* (**ECCV**), 2022 [Paper | Code+Data | Poster | Slides | Video]
- **Pramanick S.***, Roy A.*, Patel V., "Multimodal Learning using Optimal Transport for Sarcasm and Humor Detection", *Winter Conference on Applications of Computer Vision* (**WACV**), 2022. [Paper]
- **Pramanick S.***, Sharma S*., Dimitrov D., Aktar S., Nakov P., Chakraborty T., "MOMENTA: A Multimodal Framework for Detecting Harmful Memes and Their Targets", *Findings of Empirical Methods in Natural Language Processing* (**EMNLP**), Nov. 2021. [Paper | Code+Data | Poster | Slides]
- **Pramanick S.**, Dimitrov D., Mukherjee R., Sharma S., Aktar S., Nakov P., Chakraborty T., "Detecting Harmful Memes and Their Targets", *Findings of Annual Meeting of the Association for Computational Linguistics* (**ACL**), Aug. 2021. [Paper | Code+Data | Slides | Video]
- **Pramanick S.**, Aktar S., Chakraborty T., "Exercise? I thought you said 'Extra Fries': Leveraging Sentence Demarcations and Multi-hop Attention for Meme Affect Analysis", *AAAI Conference on Web and Social Media* (**ICWSM**), Jun. 2021. [Paper | Code | Poster | Slides | Video]

Journals

- Rakshit A., **Pramanick S.**, Bagchi A., Bhattacharyya S., "Autonomous grasping of 3-D objects by a vision-actuated robot arm using Brain–Computer Interface", **Biomedical Signal Processing and Control** [IF - 5.076], Elsevier, Feb. 2023. [Paper]
- Atri Y.*, **Pramanick S.***, Goyal V., Chakraborty T., "See, Hear, Read: Leveraging Multimodality with Guided Attention for Abstractive Text Summarization", ***Knowledge-Based Systems*** [IF - 8.139], Elsevier, Sept. 2021. [Paper | Code+Data]

Selected Honors & Awards

Jan 2021	JHU ECE Departmental Fellowship, *awarded to outstanding incoming PhD students.*
May 2019	Mitacs Globalink Research Internship, *awarded to top-ranked applicants from 15 different countries to participate in a 12-week research internship in Canadian universities.*
Oct 2016	JBNSTS Senior Scholarship, *4-year scholarship for academic excellence during B.E.*
Jan 2015	Regional Mathematical Olympiad (RMO), *ranked among top-10 students in the state.*

Technical Skills

Programming Languages: Python, MATLAB, C/C++, Mathematica, LaTeX
Libraries & Tools: Pytorch, Keras, Tensorflow

Voluntary Services

Reviewer for CVPR, ECCV, ICCV, WACV, EMNLP, ACL, TPAMI, TNNLS, TAI, TAFFC.

References

Rama Chellappa, Bloomberg Distinguished Professor, Johns Hopkins University
Pengchuan Zhang, Research Scientist, Meta AI
Li Jing, Research Scientist, OpenAI